

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 16, 2002, 16:10:20 ; Search time 16.06 Seconds
(without alignments)
746.760 Million cell updates/sec

Title: US-09-782-390-1
Perfect score: 2568
Sequence: 1 MEASRCRLSPSGDSVFNHEM.....HPWRQLPQSLVGPDLXLM 491

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTCUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfill1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2566	99.9	491	4	US-08-812-824-1
2	2109	82.1	442	3	US-09-032-365A-17
3	1484	57.8	390	4	US-08-977-865-2
4	1204.5	46.9	518	3	US-09-032-365A-58
5	1203.5	46.9	512	3	US-09-032-365A-60
6	1196.5	46.6	506	1	US-08-631-200-8
7	1196.5	46.6	506	2	US-08-829-553-8
8	1196.5	46.6	506	2	US-08-922-267A-8
9	1196.5	46.6	506	2	US-08-936-707A-8
10	1196.5	46.6	506	2	US-08-936-706A-8
11	1196.5	46.6	506	3	US-09-032-365A-62
12	1196.5	46.6	506	3	US-09-032-365A-62
13	1196.5	46.6	506	4	US-08-812-824-4
14	1196.5	46.6	506	4	US-09-406-071-8
15	1196.5	46.6	506	4	US-08-955-918C-10
16	1192.5	46.4	561	3	US-08-714-991-27
17	1192.5	46.4	561	3	US-09-032-365A-10
18	1186	46.2	505	1	US-08-631-200-2
19	1186	46.2	505	1	US-08-630-592-4
20	1186	46.2	505	1	US-08-714-991-4
21	1186	46.2	505	2	US-08-829-553-2
22	1186	46.2	505	2	US-08-922-267A-2
23	1186	46.2	505	2	US-08-936-707A-2
24	1186	46.2	505	2	US-08-936-706A-2
25	1186	46.2	505	3	US-09-248-203-2
26	1186	46.2	505	3	US-09-032-365A-4
27	1186	46.2	505	4	US-08-812-824-3

28	1186	46.2	505	4	US-09-406-071-2	Sequence 2, Appli
29	1167.5	45.5	504	4	US-08-955-918C-7	Sequence 7, Appli
30	1105.5	43.0	460	1	US-08-630-592-7	Sequence 7, Appli
31	1105.5	43.0	460	1	US-08-714-991-7	Sequence 7, Appli
32	1105.5	43.0	460	1	US-09-032-365A-8	Sequence 8, Appli
33	1104	43.0	442	3	US-08-977-865-4	Sequence 4, Appli
34	1095	42.6	459	1	US-08-630-592-2	Sequence 2, Appli
35	1095	42.6	459	1	US-08-714-991-2	Sequence 2, Appli
36	1095	42.6	459	1	US-09-032-365A-2	Sequence 2, Appli
37	931	36.3	542	1	US-08-701-380-2	Sequence 2, Appli
38	931	36.3	542	3	US-09-032-365A-13	Sequence 13, Appli
39	920	35.8	349	4	US-08-955-918C-2	Sequence 2, Appli
40	915.5	35.7	285	1	US-08-631-200-15	Sequence 15, Appli
41	915.5	35.7	285	2	US-08-829-553-15	Sequence 15, Appli
42	915.5	35.7	285	2	US-08-922-267A-15	Sequence 15, Appli
43	915.5	35.7	285	2	US-08-936-707A-15	Sequence 15, Appli
44	915.5	35.7	285	2	US-08-936-706A-15	Sequence 15, Appli
45	915.5	35.7	285	3	US-09-248-203-15	Sequence 15, Appli

ALIGNMENTS

RESULT 1
US-08-812-824-1
; Sequence 1, Application US/08812824
; Patent No. 6204372
GENERAL INFORMATION:
; APPLICANT: Labrie, Samuel T.
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: NOVEL TUBBY HOMOLOGUE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/812,824
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
; LENGTH: 491 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: HNT2MOT01
; CLONE: Consensus
US-08-812-824-1

Query Match 99.9%; Score 2566; DB 4; Length 491;
Best Local Similarity 100.0%; Pred. No. 2.5e-253;

Matches 491; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEASRCRLSPSGDSVVFHEEMKMKROAKLDYORLLLEKROKKRLLEPPMOPNEARLRA 60
 1 MEASRCRLSPSGDSVVFHEEMKMKROAKLDYORLLLEKROKKRLLEPPMOPNEARLRA 60

QY 61 KPRASDEQTPLVNCHTSPHSNVILHGIDGPAAYLKPDVEHAPSVSSSVEEDAENTVDAS 120
 61 KPRASDEQTPLVNCHTSPHSNVILHGIDGPAAYLKPDVEHAPSVSSSVEEDAENTVDAS 120

QY 121 KPGLOERLOKHDISESYNFDEETDGISQSACLERPNSASSONSTDGTGSSATAAPADN 180
 121 KPGLOERLOKHDISESYNFDEETDGISQSACLERPNSASSONSTDGTGSSATAAPADN 180

QY 181 LIGDIDLEDFVYSPAOGVTVRCRIIRDKRGMDRGLEFPTYWYLEKEENOKIFLLAARK 240
 181 LIGDIDLEDFVYSPAOGVTVRCRIIRDKRGMDRGLEFPTYWYLEKEENOKIFLLAARK 240

QY 241 RKSKTANLYLISIDPVDLSREGESEYVGLRSNLMGKFTVYDRGICPMKGRGLVGAHTR 300
 241 RKSKTANLYLISIDPVDLSREGESEYVGLRSNLMGKFTVYDRGICPMKGRGLVGAHTR 300

QY 301 QELAAISYETNVILGFKGPKMSYIIPGMTLNHKQIPYOPONNHDSLSRMONTMENLVE 360
 301 QELAAISYETNVILGFKGPKMSYIIPGMTLNHKQIPYOPONNHDSLSRMONTMENLVE 360

QY 361 LHNKAPVWNSDTQSYVLFNFRGRTQASVKNFQIVHKNDPDYIVMOFGRAADVFTLDYNY 420
 361 LHNKAPVWNSDTQSYVLFNFRGRTQASVKNFQIVHKNDPDYIVMOFGRAADVFTLDYNY 420

QY 421 PLCAVQAFGIGLSSFDKRTQTLRMOELCELHGHSAASLVHRTACORWGHFWHQLOPOS 480
 421 PLCAVQAFGIGLSSFDKRTQTLRMOELCELHGHSAASLVHRTACORWGHFWHQLOPOS 480

QY 481 SLVGPDLXLHM 491
 481 SLVGPDLXLHM 491

RESULT 2

US-09-032-365A-17
 ; Sequence 17, Application US/09032365A
 ; Patent No. 6114502

GENERAL INFORMATION:

APPLICANT: NO. 6114502ch, Michael
 APPLICANT: Nishina, Patsy
 APPLICANT: Naggart, Juergen
 APPLICANT: No. 6114502en-Trauth, Konrad
 TITLE OF INVENTION: GENE FAMILY ASSOCIATED WITH
 TITLE OF INVENTION: NEUROSENSORY DEFECTS
 NUMBER OF SEQUENCES: 67

CORRESPONDENCE ADDRESS:

ADDRESSEE: Bozicevic & Reed, LLP
 STREET: 285 Hamilton Avenue, Suite 200
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94301

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION NUMBER: US/09/032.365A

FILING DATE:

CLASSIFICATION: 536
 PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Sherwood, Pamela J
 REGISTRATION NUMBER: 36,677

REFERENCE/DOCKET NUMBER: SEQ-2CIP2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-327-3400
 TELEFAX: 650 327-3231

TELEX:

INFORMATION FOR SEQ ID NO: 17:

SEQUENCE CHARACTERISTICS:

LENGTH: 442 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-032-365A-17

Query Match

Best Local Similarity 93.4%; Score 2109; DB 3; Length 442;

Matches 410; Conservative 3; Mismatches 26; Indels 0; Gaps 0;

QY 1 MEASRCRLSPSGDSVVFHEEMKMKROAKLDYORLLLEKROKKRLLEPPMOPNEARLRA 60
 1 MEASRCRLSPSGDSVVFHEEMKMKROAKLDYORLLLEKROKKRLLEPPMOPNEARLRA 60

QY 61 KPRASDEQTPLVNCHTSPHSNVILHGIDGPAAYLKPDVEHAPSVSSSVEEDAENTVDAS 120
 61 KPRASDEQTPLVNCHTSPHSNVILHGIDGPAAYLKPDVEHAPSVSSSVEEDAENTVDAS 120

QY 121 KPGLOERLOKHDISESYNFDEETDGISQSACLERPNSASSONSTDGTGSSATAAPADN 180
 121 KPGLOERLOKHDISESYNFDEETDGISQSACLERPNSASSONSTDGTGSSATAAPADN 180

QY 181 LIGDIDLEDFVYSPAOGVTVRCRIIRDKRGMDRGLEFPTYWYLEKEENOKIFLLAARK 240
 181 LIGDIDLEDFVYSPAOGVTVRCRIIRDKRGMDRGLEFPTYWYLEKEENOKIFLLAARK 240

QY 241 RKSKTANLYLISIDPVDLSREGESEYVGLRSNLMGKFTVYDRGICPMKGRGLVGAHTR 300
 241 RKSKTANLYLISIDPVDLSREGESEYVGLRSNLMGKFTVYDRGICPMKGRGLVGAHTR 300

QY 301 QELAAISYETNVILGFKGPKMSYIIPGMTLNHKQIPYOPONNHDSLSRMONTMENLVE 360
 301 QELAAISYETNVILGFKGPKMSYIIPGMTLNHKQIPYOPONNHDSLSRMONTMENLVE 360

QY 361 LHNKAPVWNSDTQSYVLFNFRGRTQASVKNFQIVHKNDPDYIVMOFGRAADVFTLDYNY 420
 361 LHNKAPVWNSDTQSYVLFNFRGRTQASVKNFQIVHKNDPDYIVMOFGRAADVFTLDYNY 420

QY 421 PLCAVQAFGIGLSSFDKRTQTLRMOELCELHGHSAASLVHRTACORWGHFWHQLOPOS 480
 421 PLCAVQAFGIGLSSFDKRTQTLRMOELCELHGHSAASLVHRTACORWGHFWHQLOPOS 480

RESULT 3

US-08-977-865-2

; Sequence 2, Application US/08977865
 ; Patent No. 6187908

GENERAL INFORMATION:

APPLICANT: Terrell, Jonathan A.
 APPLICANT: Testa, Tania T.
 APPLICANT: Glozier, Israel S.
 APPLICANT: Hughes, Stephen A.
 APPLICANT: Doe, Trudy R.
 TITLE OF INVENTION: No. 6187908el Compounds
 NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ratner & Prestia
 STREET: P.O. Box 980
 CITY: Valley Forge
 STATE: PA
 COUNTRY: USA
 ZIP: 19482

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/977, 865
FILING DATE: Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 96 24433.0
FILING DATE: No. 6187908ember 25, 1996
FILING DATE: and
APPLICATION NUMBER: EP 97 307877.7
FILING DATE: October 6, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Prestia, Paul F.
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH30671
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 390 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-977-865-2

Query Match 57.8%; Score 1484; DB 4; Length 390;
Best Local Similarity 97.7%; Pred. No. 5.4e-143;
Matches 291; Conservative 1; Mismatches 4; Indels 2; Gaps 1;
QY 21 MKMRQAKLDYORLLLEKRRKRLPEFVQNPBARLRANPRASDBQTPLVNCHTPHSN 80
DB 1 MKMRQAKLDYORLLLEKRRKRLPEFVQNPBARLRANPRASDBQTPLVNCHTPHSN 60
QY 81 VILHGIDGPAVLPKDEYHAPSVSSVVEEDAEVNTVTSKPGIQRKQKHDSVNF 140
DB 61 VILHGIDGPAVLPKDEYHAPSVSSVVEEDAEVNTVTSKPGIQRKQKHDSVNF 120
QY 141 EETGIGISACLERPNSASONSSTDTGSGSATAPADNLGIDIDDEDVYSPAGOV 200
DB 121 EETGIGISACLERPNSASONSSTDTGSGSATAPADNLGIDIDDEDVYSPAGOV 180
QY 201 TVCRILIRDKRGMDRGLEPTTYMYLLEKENQIFLAAKRRKSKTANTYLISIDPVL 260
DB 181 TVCRILIRDKRGMDRGLEPTTYMYLLEKENQIFLAAKRRKSKTANTYLISIDPVL 240
QY 261 EGESVYGLRSLNMGTKTYYDRGICPMKRGVGAANTROELALISTETVNLGFGP 318
DB 241 EGESVYGLRSLNMGTKTYYDRGICPMKRGVGAANTROELALISTETVNLGFGP 296

RESULT 4
US-09-032-365A-58
Sequence 58, Application US/09032365A
Patent No. 6114502
GENERAL INFORMATION:
APPLICANT: No. 6114502th, Michael
APPLICANT: Nishina, Patsy
APPLICANT: Nagart, Juergen
APPLICANT: No. 6114502en-Trauth, Konrad
TITLE OF INVENTION: GENE FAMILY ASSOCIATED WITH
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Bozicevic & Reed, LLP
STREET: 285 Hamilton Avenue, Suite 200
CITY: Palo Alto
STATE: CA

COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/032, 365A
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J.
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: SEQ-2CIP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-327-3400
TELEFAX: 650 327-3231
TELEX:
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 518 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-032-365A-58

Query Match 46.9%; Score 1204.5; DB 3; Length 518;
Best Local Similarity 49.2%; Pred. No. 3.1e-114;
Matches 250; Conservative 74; Mismatches 111; Indels 73; Gaps 10;
QY 2 EASRC--RLSPGSDSYFHEEMKMRQAKLDYORLLLEKRRKRLPEFVQNPBARLR 59
DB 11 EKSCCARQTSRLGSHVLDGRLRQKIDRQALLLEQKRRQPLVQVANNADGRPS 70
QY 60 AKPRASDEQTPLVNCHTPHSNVLH----- 84
DB 71 RRAROSEQAPLVESYLSGSGTSYOVQEADSLASVOLGATRPAPASAKRTKAATAG 130
QY 85 -----GIDGPAVLPK-DEYHAPSVSSVVEED-AENTVDTASAPG----- 123
DB 131 QGGARREKKGKRGKSGPALAEKSEAGPVOILFVGSDHAQDAGETAGGGERPSG 190
QY 124 --LOERLOKHDSVNFDE---ETDGISOSACLE--RPNSSASONSSTDTGSGSATA 174
DB 191 QDLRATMQRKGISSMSFDEDEDEDESSSSQLSNTRPSATSRKSVREASAPSP 250
QY 175 A--QPADNLGIDIDDEDVYSPAGOVYKCIITDKRGMDGLEPTTYMYLLEKENQ 232
DB 251 APEQPVDP--VEYODLEEFALRPAPQGITIKRITRDKKMDGMVPTFLIHDRDGRK 307
QY 233 IFLAARRRKSKTANTYLISIDPVLDSREGESVYGLRSLNMGTKTYYDRGICPMKRG 292
DB 308 VFLAGKRRKSKTANTYLISIDPVLDSREGDSITGKLRSLNMGTKTYYDRGICPMK 367
QY 293 LVGAANT-ROELALISTETVNLGFGKPRKMSVILPGNTLNHKOIPYOPONNHDLSLR 351
DB 368 STLESGRLRQELAAVCEYETVNLGFGKPRKMSVILPGNNVHERVSI RPRNEHETLLAR 427
QY 352 NRTMENLVEIHNKAPVANSSTQSYVNLFRKRVYQASVKNPQYIHNKDPDPIYVQFGRV 411
DB 428 NKTESIIEIQNTPVWNDTQSYVNLFRGRVYQASVKNPQYIHNKDPDPIYVQFGRV 487
QY 412 DVFTIDYNYPLCAVOAAGFGLISSEFDKRI 439
DB 488 DVFTIDYNYPLCALQAFALISSEFDKRI 515

```

RESULT 5
US-09-032-365A-60
; Sequence 60, Application US/09032365A
; Patent No. 6114502
; GENERAL INFORMATION:
; APPLICANT: No. 6114502ch, Michael
; APPLICANT: Nishina, Patsy
; APPLICANT: Nagart, Juergen
; APPLICANT: No. 6114502en-Trauth, Konrad
; TITLE OF INVENTION: GENE FAMILY ASSOCIATED WITH
; TITLE OF INVENTION: NEUROSENSORY DEFECTS
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/032, 365A
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: SEQ-2C1P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400
; TELEFAX: 650-327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 512 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-032-365A-60

Query Match 46.9%; Score 1203.5; DB 3; Length 512;
Best Local Similarity 49.7%; Pred. No. 3,8e-114;
Matches 247; Conservative 73; Mismatches 106; Indels 71; Gaps 9;

QY 11 SGDSVFHEEMKMKROAKLDYORLLERKORRKLPEFVYQNPPEARLRAPRASDEQTP 70
DB 16 SYDSVLDEGRNLROQKLDROPRALLEOKKROEPLMWQANAGRRSRAROSEQAP 75
QY 71 LVNCHTPHSNVILH----- 84
DB 76 LVESTLSSSGSTSYQVQADSLASVQIGATRPAPASAKRTKAATAGGCGAARKEKKG 135
QY 85 ---GIDGPAAYLK-PDEVHAPSVSSVVEED-AENTVDTASKPG-----LQERLOKHD 132
DB 136 KHKGTSGPAALAEKSEAGQVQILTVQSDHADAGETAGGGERPSGQDLRATWQRKG 195
QY 133 ISESYNPE---ETDGISOSACLE---RPNASASSONTDTCTGSGSATAA--QPADNLG 183
DB 136 ISSNSGFDEDEDEENSSQSLNSNTSPSSATSRKSVREAPASAPSTAEQPDV--V 252
QY 184 DIDLDEEVYSPAPQVTVRCRIIRDKRGMRLGFTYYATLEKEENKIFLLAARKKK 243
DB 253 EVQDLEEFALRRAPGQITIKCRITRDKKGMKGMWPTFTFLRLDRDGGKVVLAGRRKKK 312

```

```

QY 244 SKTANYLISIDPVDLSREGSEYVGKLRSLNMGKTFETYDGRICPMKRGVLGAHT--ROE 302
DB 313 SKTSNYLISVDPDLSKGGDSYIGKLRSLNMGKTFETYDGVNPDQASSSTLESGLTROE 372
QY 303 LAALSYETNLVGFGRPKRMSVILPGMTLNHKOIPIYOPQNNHDSILSRMQRMENTVELH 362
DB 373 LAACVETNLVGFGRPKRMSVILPGMTLNHKOIPIYOPQNNHDSILSRMQRMENTVELH 432
QY 363 NKAIVMSDQSYLVNLRGRVTAQSVYKRNQIVKKNDDIYVQGRADVFTLDNYPL 422
DB 433 NKTPIVMDQSYLVNLRGRVTAQSVYKRNQIVKKNDDIYVQGRADVFTLDNYPL 492
QY 423 CAVQAFGIGLSSFDKRI 439
DB 493 CALQAFALISLSPSKL 509

RESULT 6
US-08-631-200-8
; Sequence 8, Application US/08631200
; Patent No. 5646040
; GENERAL INFORMATION:
; APPLICANT: Kleyn, Patrick W.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennle & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/631, 200
; FILING DATE: 12-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-057
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 506 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-631-200-8

Query Match 46.6%; Score 1196.5; DB 1; Length 506;
Best Local Similarity 49.6%; Pred. No. 1.9e-113;
Matches 245; Conservative 73; Mismatches 105; Indels 71; Gaps 9;

QY 14 SYFHEEMKMKROAKLDYORLLERKORRKLPEFVYQNPPEARLRAPRASDEQTP 73
DB 13 SYLDDSGRNLROQKLDROPRALLEOKKROEPLMWQANADGRRRAROSEQAPLVE 72
QY 74 CHTPHSNVILH----- 84
DB 73 SYLSSSGSTSYQVQADSLASVQIGATRPAPASAKRTKAATAGGCGAARKEKKGK 132
QY 85 GIDGPAAYLK-PDEVHAPSVSSVVEED-AENTVDTASKPG-----LQERLOKHDISE 135

```

Db 133 GTSGPAALAEKSEAGQVQILTVGOSDHAQAGETAAGGERPSGODLRATMQRKGISS 192
 QY 136 SVNDE---ETDGISQACLE---RPNASSONSTDTGTSATFAA--QPADNLGDID 186
 Db 193 SMSFDEDEDEENSSSSQSLNSNTRPSSATSRKSVREAAAPSPAPQVD--VEVQ 249
 QY 187 DLEDFVSPAPQGVYVRCRIIRDKRGMDRGLFTTYMYLKEENOKIFLLAARKRKSKT 246
 Db 250 DLEEFALRPAPQGITIKRITRDKKGMVPTFYELHLDREDEGKVFLLAGRRKRSKT 309
 QY 247 ANYLISIDPDLRSREGESYVGLRSNLMGKFTYVDRGICPMGRGLVGAANT-ROELAA 305
 Db 310 SNYLSVDPDLRSRGDSYIGKLRSLNMGKFTYVNGVNPQKASSSTLESGLRQELAA 369
 QY 306 ISEYTNVLFGRKPRKMSVTPGMNVHERVSIIRPNEHETLLARWONKNTESIELONKT 429
 Db 370 VCETNVLFGRKPRKMSVTPGMNVHERVSIIRPNEHETLLARWONKNTESIELONKT 429
 QY 366 PWMNSDTQSVYLNFRGVTQASVKNFOIVHKNDPDYVMOFGVADVFTLDYNYPLCAV 425
 Db 430 PWMNDTQSVYLNFRGVTQASVKNFOIVHKNDPDYVMOFGVADVFTLDYNYPLCAL 489
 QY 426 QAFGIGLSSFDKRI 439
 Db 490 QAFALALSSFDKRI 503

RESULT 7
 US-08-829-553-8
 ; Sequence 8, Application US/08829553
 ; Patent No. 5817762

GENERAL INFORMATION:

APPLICANT: Kleyn, Patrick W.
 APPLICANT: Moore, Karen J.
 TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
 NUMBER OF SEQUENCES: 59
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Pennie & Edmonds
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/829,553
 FILING DATE: 28-MAR-1997
 CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/631,200
 FILING DATE: 12-APR-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Coruzzi, Laura A.
 REGISTRATION NUMBER: 30,742
 REFERENCE/DOCKET NUMBER: 7853-057

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-9741/8864
 TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:
 LENGTH: 506 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 IS-08-829-553-8

Query Match 46.6%; Score 1196.5; DB 2; Length 506;
 Best Local Similarity 49.6%; Pred. No. 1,9e-113;
 Matches 245; Conservative 73; Mismatches 105; Indels 71; Gaps 9;

QY 14 SVEHEEMKMRQAKLDYORLLERKORRKLREFFVQNPPEARLRARAKPRASDEQTPLVN 73
 Db 13 SVLDEGRNLROQKLDORALLEROKKROEPLMVOANADGPRRRAROSEQAPLVE 72
 QY 74 CHPHSVILH----- 84
 Db 73 SYLSSSGTSYQVQVADSLASVOLGATRPAPASAKRTAAATVAGGGAARKEKKKKH 132
 QY 85 GICGPAVILK-PDEVHAPSVSSVVEED-AENTVDTASKP-----LOERLOKHIDE 135
 Db 133 GTSGPAALAEKSEAGQVQILTVGOSDHAQAGETAAGGERPSGODLRATMQRKGISS 192
 QY 136 SVNDE---ETDGISQACLE---RPNASSONSTDTGTSATFAA--QPADNLGDID 186
 Db 193 SMSFDEDEDEENSSSSQSLNSNTRPSSATSRKSVREAAAPSPAPQVD--VEVQ 249
 QY 187 DLEDFVSPAPQGVYVRCRIIRDKRGMDRGLFTTYMYLKEENOKIFLLAARKRKSKT 246
 Db 250 DLEEFALRPAPQGITIKRITRDKKGMVPTFYELHLDREDEGKVFLLAGRRKRSKT 309
 QY 247 ANYLISIDPDLRSREGESYVGLRSNLMGKFTYVDRGICPMGRGLVGAANT-ROELAA 305
 Db 310 SNYLSVDPDLRSRGDSYIGKLRSLNMGKFTYVNGVNPQKASSSTLESGLRQELAA 369
 QY 306 ISEYTNVLFGRKPRKMSVTPGMNVHERVSIIRPNEHETLLARWONKNTESIELONKT 429
 Db 370 VCETNVLFGRKPRKMSVTPGMNVHERVSIIRPNEHETLLARWONKNTESIELONKT 429
 QY 366 PWMNSDTQSVYLNFRGVTQASVKNFOIVHKNDPDYVMOFGVADVFTLDYNYPLCAV 425
 Db 430 PWMNDTQSVYLNFRGVTQASVKNFOIVHKNDPDYVMOFGVADVFTLDYNYPLCAL 489
 QY 426 QAFGIGLSSFDKRI 439
 Db 490 QAFALALSSFDKRI 503

RESULT 8

US-08-922-267A-8
 ; Sequence 8, Application US/08922267A
 ; Patent No. 5861239

GENERAL INFORMATION:

APPLICANT: Kleyn, Patrick W.
 APPLICANT: Moore, Karen J.
 TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
 NUMBER OF SEQUENCES: 82
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Pennie & Edmonds LLP
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/922,267A
 FILING DATE: 2-SEP-1997
 CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/829,553
 FILING DATE: 28-MAR-1997
 CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/631,200
 FILING DATE: 12-APR-1996
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: Coruzzi, Laura A.
 REGISTRATION NUMBER: 30,742
 REFERENCE/DOCKET NUMBER: 7853-085
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-9741/8864
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 506 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 US-08-922-267A-8

Query Match

Best Local Similarity 46.6%; Score 1196.5; DB 2; Length 506;
 Matches 245; Conservative 73; Mismatches 105; Indels 71; Gaps 9;

QY 14 SVFHEEMMMKQAKLDYQRLLEKRRKRLPEFMVQPNPEARLRARPRASDEQTPLVN 73
 DB 13 SYLDEGRNLROOKLDROKRLLEOKKKRQEPPLMVQANADGPRSRARQSEQAPLVE 72
 QY 74 CHTPHSNVLH----- 84
 DB 73 SYLSSSGSTSYQVEADSLASVOLGATRPAPASAKRTKAATAGGGAARKEKKKKH 132
 QY 85 GIDPAVAVLK-PDEVHAPSVSSVVEED-AENTVDTASKPG-----LOERLOKHDISE 135
 DB 133 CTSGRPALAEDKSEAQGPVQIITVQOSDHAQAGETAAGGGERPSGODLRATYQKRGISS 192
 QY 136 SVNFEDE---ETDGISQSAACLE--RPNSSASSONSTDGTGSGSATPA--QPADNLLGDI 186
 DB 193 SMSFDEDEDEEENSSSSQSLNSTRPSATSRSKSVREASAPSPAPAEQVVD---VEVQ 249
 QY 187 DLEDFVSPAPQGVTVRCRIIRDKRGMDRGLPTTYWYLEKEENKIFLLAARKKRSKT 246
 DB 250 DLEEFALRPAPQGITKICITRDKKGMDBMYPTFLHLDRDEGKVFLLAGRRKKRSKT 309
 QY 247 ANYLISIDPVDLSREGESYVGLKRSNLMTGKFTYVDRGICPMKGRGLVGAHT-ROELAA 305
 DB 310 SNYLSVDPDTLSRGSDSYIGKLRSLNMGKFTYVDNGVNPQKASSSTLESGLTLOELAA 369
 QY 306 ISEFTNVLGFKGPRKMSVILPGMTLNHKOIPYQPNNHDSLSRWQNTMENTVELANKA 365
 DB 370 VCETNVLGFKGPRKMSVILPGMNVHERVSIIRPNEHETLLARWQNKTESITIELQNT 429
 QY 366 PWNMSDQSYVNLFRGVTYQASVKNFOIYVKNDRPYIWMQFGRVADVFTLDYNTPLCAV 425
 DB 430 PWNMDTQSYVNLFRGVTYQASVKNFOIIGHNDPDYIWMQFGRVADVFTMDYNTPLCAL 489
 QY 426 QAFGIGLSFDPKRI 439
 DB 490 QAFATLALSFDKSL 503

RESULT 9

US-08-936-707A-8

Sequence 8, Application US/08936707A
 Patent No. 5871931

GENERAL INFORMATION:
 APPLICANT: Kieyn, Patrick W.
 APPLICANT: Moore, Karen J.
 TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
 NUMBER OF SEQUENCES: 60
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennie & Edmonds LLP

STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036-2711
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/936,707A
 FILING DATE: 24-SEP-1997
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Coruzzi, Laura A.
 REGISTRATION NUMBER: 30,742
 REFERENCE/DOCKET NUMBER: 7853-100
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-9741/8864
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 506 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 US-08-936-707A-8

Query Match

Best Local Similarity 46.6%; Score 1196.5; DB 2; Length 506;
 Matches 245; Conservative 73; Mismatches 105; Indels 71; Gaps 9;

QY 14 SVFHEEMMMKQAKLDYQRLLEKRRKRLPEFMVQPNPEARLRARPRASDEQTPLVN 73
 DB 13 SYLDEGRNLROOKLDROKRLLEOKKKRQEPPLMVQANADGPRSRARQSEQAPLVE 72
 QY 74 CHTPHSNVLH----- 84
 DB 73 SYLSSSGSTSYQVEADSLASVOLGATRPAPASAKRTKAATAGGGAARKEKKKKH 132
 QY 85 GIDPAVAVLK-PDEVHAPSVSSVVEED-AENTVDTASKPG-----LOERLOKHDISE 135
 DB 133 CTSGRPALAEDKSEAQGPVQIITVQOSDHAQAGETAAGGGERPSGODLRATYQKRGISS 192
 QY 136 SVNFEDE---ETDGISQSAACLE--RPNSSASSONSTDGTGSGSATPA--QPADNLLGDI 186
 DB 193 SMSFDEDEDEEENSSSSQSLNSTRPSATSRSKSVREASAPSPAPAEQVVD---VEVQ 249
 QY 187 DLEDFVSPAPQGVTVRCRIIRDKRGMDRGLPTTYWYLEKEENKIFLLAARKKRSKT 246
 DB 250 DLEEFALRPAPQGITKICITRDKKGMDBMYPTFLHLDRDEGKVFLLAGRRKKRSKT 309
 QY 247 ANYLISIDPVDLSREGESYVGLKRSNLMTGKFTYVDRGICPMKGRGLVGAHT-ROELAA 305
 DB 310 SNYLSVDPDTLSRGSDSYIGKLRSLNMGKFTYVDNGVNPQKASSSTLESGLTLOELAA 369
 QY 306 ISEFTNVLGFKGPRKMSVILPGMTLNHKOIPYQPNNHDSLSRWQNTMENTVELANKA 365
 DB 370 VCETNVLGFKGPRKMSVILPGMNVHERVSIIRPNEHETLLARWQNKTESITIELQNT 429
 QY 366 PWNMSDQSYVNLFRGVTYQASVKNFOIYVKNDRPYIWMQFGRVADVFTLDYNTPLCAV 425
 DB 430 PWNMDTQSYVNLFRGVTYQASVKNFOIIGHNDPDYIWMQFGRVADVFTMDYNTPLCAL 489
 QY 426 QAFGIGLSFDPKRI 439
 DB 490 QAFATLALSFDKSL 503

RESULT 10

US-08-936-706A-8
Sequence 8, Application US/08936706A
Patent No. 5876919
GENERAL INFORMATION:
APPLICANT: Kleyn, Patrick W.
APPLICANT: Moore, Karen J.
TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennle & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936,706A
FILING DATE: 24-SEP-1997
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-099
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 506 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-936-706A-8

Query Match 46.6%; Score 1196.5; DB 2; Length 506;
Best Local Similarity 49.6%; Pred. No. 1.9e-113;
Matches 245; Conservative 73; Mismatches 105; Indels 71; Gaps 9;
QY 14 SVFHEEMKMKROAKLDYQRLLEKRRKRLPEPMVQPNPEARLRRAPRASDEQTPLYN 73
DB 13 SVLDEGRNLROQKLDROKRLLEKRRKRLPEPMVQPNPEARLRRAPRASDEQTPLYN 72
QY 74 CHTPNSVILH----- 84
DB 73 SYLSSSGSTSYOVQADSLASVOLGATPTAPASAKRTKAATAGGGAARKEKKGNH 132
QY 85 GIDGPAAYLK-PDEVHAPSVSSSVYED-AENTVDTASKPG-----LOERLOKHDISE 135
DB 133 GTSGRNALAEKESPAQGVQVLTITVQSDHQAODAGETAGGGERPBGODLRATMQRKGISS 192
QY 136 SVNDE---ETDGISASACLE--RPNASASSNSTDTGTSGSATAA--QPADNLLGDI 186
DB 193 SMSFDEDEDEDESSSSQLNSNTRPSSATSRKSVRAASAPSPTABEQPVD--VEVQ 249
QY 187 DLEDEVSAPAGVVRRIIFDRKGMGRGLFPTTYMTELEKEMNKITLLAARKKKKKT 246
DB 250 DLEELRLRPAPOGITIKCTTRDKGMGRMPTVFLHLDKDGKVFLLGRKKKKKT 309
QY 247 ANYLISIDPVDLSREGESYVGLKRLNMGTKFTVYDRGICPMKGRGLGAANT-ROELAA 305
DB 310 SNYLLSVDPDTLDSRGDSYIGKRLSNLMGTFTVYDNGVNPQKASSSTLSESTLQELAA 369
QY 306 ISEYTNVLGFKPRKMSVITPMTLNHKOIPYOPONNHDSLSLRQNTMENVELANKA 365
DB 370 VCETYNVLGFKPRKMSVITPMTLNHKOIPYOPONNHDSLSLRQNTMENVELANKA 429

QY 366 PWNNSDTSQSYVLNFRGHTQASVKNFQIVKNDPDYIVMFGHADVFTLDYNNPLCAV 425
DB 430 PWNNDPTQSYVLNFRGHTQASVKNFQIIGHNDPDYIVMFGHADVFTMDYNNPLCAL 489
QY 426 QAFGIGLSFDPKRI 439
DB 490 QAFALISFDSKL 503
RESULT 11
US-09-248-203-8
Sequence 8, Application US/09248203
Patent No. 6043346
GENERAL INFORMATION:
APPLICANT: Kleyn, Patrick W.
APPLICANT: Moore, Karen J.
TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennle & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/248,203
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/936,707
FILING DATE: 24-SEP-1997
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-100
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 506 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-09-248-203-8
Query Match 46.6%; Score 1196.5; DB 3; Length 506;
Best Local Similarity 49.6%; Pred. No. 1.9e-113;
Matches 245; Conservative 73; Mismatches 105; Indels 71; Gaps 9;
QY 14 SVFHEEMKMKROAKLDYQRLLEKRRKRLPEPMVQPNPEARLRRAPRASDEQTPLYN 73
DB 13 SVLDEGRNLROQKLDROKRLLEKRRKRLPEPMVQPNPEARLRRAPRASDEQTPLYN 72
QY 74 CHTPNSVILH----- 84
DB 73 SYLSSSGSTSYOVQADSLASVOLGATPTAPASAKRTKAATAGGGAARKEKKGNH 132
QY 85 GIDGPAAYLK-PDEVHAPSVSSSVYED-AENTVDTASKPG-----LOERLOKHDISE 135
DB 133 GTSGRNALAEKESPAQGVQVLTITVQSDHQAODAGETAGGGERPBGODLRATMQRKGISS 192
QY 136 SVNDE---ETDGISASACLE--RPNASASSNSTDTGTSGSATAA--QPADNLLGDI 186


```

TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 506 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1305497
US-08-812-824-4

Query Match          46.6%; Score 1196.5; DB 4; Length 506;
Best Local Similarity 49.6%; Pred. No. 1.9e-113;
Matches 245; Conservative 73; Mismatches 105; Indels 71; Gaps 9;

QY 14 SVFHEEMKMRQAKIDYQRLLEKROKRLLEPPMVQPNPEARLRRAKPRASDEQPLVN 73
   || : || : || : || : || : || : || : || : || : || : || : || : ||
Db 13 SYLDEGRNLROQKIDRQALLLEQKKRKQEPRLMVQAMNDRGRSRARQSEQAPLVE 72

QY 74 CHTPHSNVILH-----84
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 73 SYLSSGSTSYQVQEADSLASVOLGATRPAPASAKRTKAATAGGGGGAARKEKKGKH 132

QY 85 GIDGPAALVK-PDEVHAPSVSSVVEED-AENTVDTASKPG-----LQERLOKHIDISE 135
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 133 GTSGPALAEKSEAGCPVQILTVGQSDHAODAGETAGGGERPSGODLRATMQRKGISS 192

QY 136 SVNFDE---ETDGISQSACLE---RPNSSASQNSTDTGSGSATAA--QPADNLLGDID 186
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 193 SMSFDEDEDEEENSSSSQLNSNTRPSATSRSKSVREASAPSPAPQEPVD---VEYQ 249

QY 187 DLEDFYSPAPQGVTVRCRIIRDKRGMDRGLFPYYMYLKEENQKIFLLAARKKRSKT 246
   || : || : || : || : || : || : || : || : || : || : || : || : ||
Db 250 DLEEFALRPAPQGITCRITRDKKGMDRGMPYFYLHLDREGQKVFLLAGRRKKRSKT 309

QY 247 ANYLISIDPVDLSREGESYVGKLRSLMGTFTYVDGICPMKRGCLVGAHT-ROELAA 305
   || : || : || : || : || : || : || : || : || : || : || : || : ||
Db 310 SNLILSVDPDLSRGDSITIGKLRSLMGTFTYDNGVNPQKASSSTLESGLROELAA 369

QY 306 ISETYNVLGFKPRKMSVILPGMTLNHKOIPYQPNHNSLSRMQNTMENLVELAHKA 365
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 370 VCETYNVLGFKPRKMSVILPGMNMVHERVSIIRPNEHETLLARQNKNTESILQNK 429

QY 366 PWNNSDTOSYVLNFGRYTQASVKNFQYVHKNDPDIYVQSGRVADVFTLDYNYPLCAV 425
   || : || : || : || : || : || : || : || : || : || : || : || : ||
Db 430 PYWMDTOSYVLNFGRYTQASVKNFQYVHKNDPDIYVQSGRVADVFTLDYNYPLCAL 489

QY 426 QAFGIGLSSFDKRI 439
   || : || : || : || : || : || : || : || : || : || : || : || : ||
Db 490 QAFALALSSFDKRL 503

RESULT 14
US-09-406-071-8
Sequence 8, Application US/09406071
Patent No. 6207386
GENERAL INFORMATION:
APPLICANT: Kleya, Patrick W.
APPLICANT: Moore, Karen J.
TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
NUMBER OF SEQUENCES: 60
DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.

```

```

ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/406,071
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/936,707
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-100
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 506 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-09-406-071-8

```

```

Query Match          46.6%; Score 1196.5; DB 4; Length 506;
Best Local Similarity 49.6%; Pred. No. 1.9e-113;
Matches 245; Conservative 73; Mismatches 105; Indels 71; Gaps 9;

QY 14 SVFHEEMKMRQAKIDYQRLLEKROKRLLEPPMVQPNPEARLRRAKPRASDEQPLVN 73
   || : || : || : || : || : || : || : || : || : || : || : || : ||
Db 13 SYLDEGRNLROQKIDRQALLLEQKKRKQEPRLMVQAMNDRGRSRARQSEQAPLVE 72

QY 74 CHTPHSNVILH-----84
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 73 SYLSSGSTSYQVQEADSLASVOLGATRPAPASAKRTKAATAGGGGGAARKEKKGKH 132

QY 85 GIDGPAALVK-PDEVHAPSVSSVVEED-AENTVDTASKPG-----LQERLOKHIDISE 135
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 133 GTSGPALAEKSEAGCPVQILTVGQSDHAODAGETAGGGERPSGODLRATMQRKGISS 192

QY 136 SVNFDE---ETDGISQSACLE---RPNSSASQNSTDTGSGSATAA--QPADNLLGDID 186
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 193 SMSFDEDEDEEENSSSSQLNSNTRPSATSRSKSVREASAPSPAPQEPVD---VEYQ 249

QY 187 DLEDFYSPAPQGVTVRCRIIRDKRGMDRGLFPYYMYLKEENQKIFLLAARKKRSKT 246
   || : || : || : || : || : || : || : || : || : || : || : || : ||
Db 250 DLEEFALRPAPQGITCRITRDKKGMDRGMPYFYLHLDREGQKVFLLAGRRKKRSKT 309

QY 247 ANYLISIDPVDLSREGESYVGKLRSLMGTFTYVDGICPMKRGCLVGAHT-ROELAA 305
   || : || : || : || : || : || : || : || : || : || : || : || : ||
Db 310 SNLILSVDPDLSRGDSITIGKLRSLMGTFTYDNGVNPQKASSSTLESGLROELAA 369

QY 306 ISETYNVLGFKPRKMSVILPGMTLNHKOIPYQPNHNSLSRMQNTMENLVELAHKA 365
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 370 VCETYNVLGFKPRKMSVILPGMNMVHERVSIIRPNEHETLLARQNKNTESILQNK 429

QY 366 PWNNSDTOSYVLNFGRYTQASVKNFQYVHKNDPDIYVQSGRVADVFTLDYNYPLCAV 425
   || : || : || : || : || : || : || : || : || : || : || : || : ||
Db 430 PYWMDTOSYVLNFGRYTQASVKNFQYVHKNDPDIYVQSGRVADVFTLDYNYPLCAL 489

QY 426 QAFGIGLSSFDKRI 439
   || : || : || : || : || : || : || : || : || : || : || : || : ||
Db 490 QAFALALSSFDKRL 503

RESULT 15
US-08-955-918C-10

```

Sequence 10. Application US/08955918C
Patent No. 6268130
GENERAL INFORMATION:
APPLICANT: Kleyu, Patrick, and Moore, Karen
TITLE OF INVENTION: RP Compositions and Therapeutic and
TITLE OF INVENTION: Diagnostic Uses Therefor
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
City: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/955,918C
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/697,766
FILING DATE: 29-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Mandragoras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: NMT-007CPDV2CPA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 506 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-955-918C-10

Query Match 46.6%; Score 1196.5; DB 4; Length 506;
Best Local Similarity 49.6%; Pred. No. 1.9e-113;
Matches 245; Conservative 73; Mismatches 105; Indels 71; Gaps 9;

QY 14 SYFHEEMKMRQAKLDYQRLLEKQRKKRLPEFMVQNPPEARLRAKPRASDEQTPLVN 73
DB 13 SYLDEGRRLRQOKLDRORALLEQOKKKROEPLMVQANADGRPRRRQSEQAPLVE 72
QY 74 CTPHSNVILH----- 84
DB 73 SYLSSSGSTSYVOEADSLASVQLATRPAPASAKRTKAATAGOGGAARKEKKGKH 132
QY 85 GIDGPAAYLK-PDEVHAPSVSSVVEED-AENTYVDFASKPG-----LQERLQKHISE 135
DB 133 GTSQPAALAEDEKSEAGQEPQILTVGQSDHQAQAGETAGGGERPSGODLRATMRKGISS 192
QY 136 SYNPE-----ETDGISQACLE---RPNSSASQNSTDTGTSGSATAA--OPADNLLGDID 186
DB 193 SMSFDEDEDEENSSSSQSLNSNTRPSSATSKSVREASAPSPPTAPEQPYD--VEVQ 249
QY 187 DLEDFYVSPAGVTVRCHIRDKRGMDRGLPFTYMYLKEENOKIFLLAARKKRSKT 246
DB 250 DLEEFALRAAPGITIKCKITRDKKGMGRGMPTFTLHIDREDGKKVFLLAGRKKRSKT 309
QY 247 ANYLISIDPVDLSREGESYVGLRSNLMGTFYVDRCICPMKGRGLVGAHT-RQELAA 305
DB 310 SNVLISVDPTDLSRGDSYIGKLRNLMGTFYVDNCGVNPQKASSSTLESGLTRQELAA 369
QY 306 ISETVNLGFKGRKMSVITIPGTLNHRQIPYQPNHDSILSRMONRTMENTVELHNKA 365
DB 370 VCETVNLGFKGRKMSVITIPGNNVHERVSIAPRNEHETLLARWONKNTESIIELOKNT 429

QY 366 PWNSSDTQSYVLNFRGRVTOASVKNFQIVHKNDPDIYWMQFGRVADVFTLDYNYPLCAV 425
DB 430 PWNNDTQSYVLNFRGRVTOASVKNFQILHGNDDPDIYWMQFGRVADVFTMDYNYPLCAL 489
QY 426 QAFGLSSFDKRI 439
DB 490 QAFALSSFDSKL 503

Search completed: July 16, 2002, 16:10:46
Job time: 26 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 16, 2002, 16:10:20 ; Search time 34.18 Seconds

(without alignments)
1595.589 Million cell updates/sec

Title: US-09-782-390-1

Perfect score: 2568
Sequence: 1 MEASRCRLSPSGDSVFHEEM.....HPWRQLPQSSLVGPDLXLM 491

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_032802:*
1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2566	99.9	491	AAW61198	Novel human tubby
2	2273	88.5	442	AAW61198	Novel human tubby
3	2109	82.1	442	AAW61198	Novel human tubby
4	2109	82.1	442	AAW61198	Novel human tubby
5	1617	63.0	460	AAW61198	Novel human tubby
6	1484	57.8	390	AAW61198	Novel human tubby
7	1204.5	46.9	518	AAW61198	Novel human tubby
8	1204.5	46.9	518	AAW61198	Novel human tubby
9	1203.5	46.9	512	AAW61198	Novel human tubby
10	1203.5	46.9	512	AAW61198	Novel human tubby
11	1196.5	46.6	506	AAW61198	Novel human tubby

12	1196.5	46.6	506	19	AAW54368
13	1196.5	46.6	506	20	AAW75451
14	1196.5	46.6	506	21	AAW26910
15	1192.5	46.4	561	18	AAW36489
16	1192.5	46.4	561	18	AAW36489
17	1187.5	46.2	506	18	AAW10729
18	1186	46.2	505	18	AAW36488
19	1186	46.2	505	18	AAW10728
20	1186	46.2	505	19	AAW54367
21	1186	46.2	505	20	AAW75450
22	1186	46.2	505	21	AAW26901
23	1105.5	43.0	460	18	AAW36488
24	1105.5	43.0	460	21	AAW26903
25	1104	42.6	459	18	AAW61160
26	1095	42.6	459	18	AAW36485
27	1095	42.6	459	21	AAW26400
28	931	36.3	542	18	AAW36490
29	931	36.3	542	21	AAW26905
30	924	36.0	349	19	AAW54365
31	915.5	35.7	285	18	AAW93805
32	915.5	35.7	285	19	AAW54366
33	915.5	35.7	285	20	AAW75452
34	882.5	34.4	443	22	AAW64394
35	826.5	32.2	506	18	AAW36494
36	826.5	32.2	506	21	AAW26908
37	790.5	30.8	520	18	AAW36491
38	790.5	30.8	520	21	AAW26906
39	428	16.7	308	21	AAW15639
40	428	16.7	308	21	AAW15638
41	428	16.7	389	21	AAW15637
42	397	15.5	400	21	AAW07736
43	396	15.4	379	21	AAW07737
44	388	15.1	342	21	AAW07738
45	260	10.1	50	22	ABW39124

ALIGNMENTS

RESULT	1	
ID	AAW61198	standard; Protein; 491 AA.
XX		
AC	AAW61198;	
XX		
DT	19-NOV-1998	(first entry)
XX		
DE	Novel human tubby homologue (NHT) protein.	
XX		
KW	HMT2 neuronal cell line; CDNA library; HMT2NOT01; Incyte clone 492199;	
KW	anorexia; cachexia; diabetes; obesity; cardiovascular disease; cancer;	
KW	autoimmune.	
XX		
OS	Homo sapiens.	
XX		
FX	Key	Location/Qualifiers
FT	Misc-difference 487	/note="encoded by CTN"
FT	Misc-difference 488	/label="unknown
FT		/note="encoded by TNA"
XX		
PN	W09839445-A1.	
XX		
PD	11-SEP-1998.	
XX		
PE	05-MAR-1998;	98MO-US04467.
XX		
PR	06-MAR-1997;	97US-0812824.
XX		
PA	(INCYTE-) INCYTE PHARM INC.	
XX		
PI	Labrie ST, Lal P, Murry LE;	

XX WPI: 1998-495849/42.
 DR N-PSDB: AAV47676.
 XX
 PT Human tubby homologue - used for treating or preventing eating disorders
 XX
 PS Claim 1, Figure 1A-E; 64pp; English.
 XX
 CC The present sequence represents a novel human tubby homologue (NHT) CC
 CC polypeptide, encoded by the NHT cDNA. This cDNA sequence was obtained by CC
 CC standard cloning and screening procedures, where it was first isolated in CC
 CC the HNT2 neuronal cell line cDNA library HNT2NOT01 in Inocyte clone CC
 CC 492199. A comparison of the NHT homologue with mouse and human tub genes CC
 CC revealed that NHT shares 49% identity with both of these genes. NHT can CC
 CC be used in the diagnosis, prevention or treatment of appetite and eating CC
 CC disorders, including anorexia, cachexia, diabetes and obesity. It can CC
 CC also be used to treat disorders associated with overeating, for example CC
 CC cardiovascular diseases, autoimmune and genetic disorders and some CC
 CC cancers.
 XX
 SQ Sequence 491 AA:

Query Match 99.9%; Score 2566; DB 19; Length 491;
 Best Local Similarity 100.0%; Pred. No. 1.5e-249;
 Matches 491; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEASRCRLSPSGDSVFHEEMMKMRQAKLDYORLLEKRRKRLPEFMVOPNEARLRA 60
 DB 1 measrcrlspsgdsvfheemmkmrqakldyqrlillekrrkrllepfmvgpnearlrra 60
 QY 61 KPRASDEQTPLVNCHTSHSNVILHGIDGPAAVLKPDVHAPSVSSVVEEDAENTVDAS 120
 DB 61 kprasdeqtplvnchthshsnvllhgidgpaavlkpdevhapsevssvveedaentvdas 120
 QY 121 KPGLOERLOKHDISESVPDEETDGISOSACLERPNASSONSTDTGTSATTAOPADN 180
 DB 121 kpglgerlqkhdisesvndeetdgisgsaclerpnasassqndtqtsataaqpahn 180
 QY 181 LAGDIDLDLDFYSPAPGCVTVRCRIIRKRCMDRGLFTYVWYLEKEENCKIFLAAK 240
 DB 181 llgididdldfyvspapgvtrcrlirrkrcmdrglftlywylekeengkflllaark 240
 QY 241 RKKSRTAANTLISIDPVDSREGBSVYGLKRSNLMGKTFTYDRGICPMKRGVLGAHTR 300
 DB 241 rkksrtanyllisidpvdlsregebsyglkrsnlmgktftfydrigicpmkrgvlgaahtr 300
 QY 301 OELAAISTEENLVGFKGPRKMSVILPGMTLNHKOIPIYOPQNNHDSLSRMQRMENIVE 360
 DB 301 qelaaisteyelnvlgfkprkmsvllpgmtlnhkqilpyqpqnhdslsrtqrmtenive 360
 QY 361 LHNKAPVWNSDQSYVLNFRGRTQASVKNFQIVHKNDPDYIVMOFGRYADVDFTLDYNY 420
 DB 361 lhnkavwnsdqsyvlnfrgvtqasvknfivhknbdpyivmfgryadvdfldyny 420
 QY 421 PLICAVOAFGIGSSFDKRIQTILMOELCELHROHNSAASLVHRTACQRMVGHWPOLPOS 480
 DB 421 plicavafgigissfdkriqtilmoelcelhrgohnsaaslvhrtacqrmvghwpolpos 480
 QY 481 SLVGPDLXLHM 491
 DB 481 slvgpdlxlhm 491

RESULT 2
 ID AAB95750 standard: Protein; 442 AA.
 XX AAB95750;
 XX
 DT 26-JUN-2001 (first entry)
 XX

DE Human protein sequence SEQ ID NO:18657.
 XX
 KW Human; primer: detection; diagnosis; antisense therapy; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PE 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 XX
 PR 27-AUG-1999; 99JP-0300253.
 XX
 PR 11-JAN-2000; 2000JP-0118776.
 XX
 PR 02-MAY-2000; 2000JP-0183767.
 XX
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI: 2001-318749/34.
 XX
 PT primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 PS Claim 8; SEQ ID 18657; 2537pp + CD ROM; English.

The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

Sequence 442 AA;

Query Match 88.5%; Score 2273; DB 22; Length 442;
 Best Local Similarity 99.3%; Pred. No. 4.6e-220;
 Matches 436; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEASRCRLSPSGDSVFHEEMMKMRQAKLDYORLLEKRRKRLPEFMVOPNEARLRA 60
 DB 1 measrcrlspsgdsvfheemmkmrqakldyqrlillekrrkrllepfmvgpnearlrra 60
 QY 61 KPRASDEQTPLVNCHTSHSNVILHGIDGPAAVLKPDVHAPSVSSVVEEDAENTVDAS 120
 DB 61 kprasdeqtplvnchthshsnvllhgidgpaavlkpdevhapsevssvveedaentvdas 120
 QY 121 KPGLOERLOKHDISESVPDEETDGISOSACLERPNASSONSTDTGTSATTAOPADN 180
 DB 121 kpglgerlqkhdisesvndeetdgisgsaclerpnasassqndtqtsataaqpahn 180

QY 181 LLGDDIDLEDFVYSPAPQGVTVRCRIIRDRKGMGRGLFPTYYMYLKEENOKIFLLAARK 240
 DB 181 llgddidledfvysspapgvttvrcrllrdkrgmgrglfptcyymylkeeenqkiflllaark 240
 QY 241 RKSKTANYLISIDPVDLSREGESYVGKLSNLMGTRKFTYDNGICPMKRGILVGAATHR 300
 DB 241 rkksktanyllisidpvdlsregesyvgklrsnlmgtrkftvydngicpmkgqrgilvgaahr 300
 QY 301 QELAAISYETNVLGFGPRKMSVITPGMTLNHKQIPYQPNNDLSLRMONTMENTLVE 360
 DB 301 qelaaissyenvlfgprkmsvllpgmtlnhkqipyqpnmhdsllstrwqntmentlve 360
 QY 361 LHNKAPVNSDTQSYVLNFRGRVTOASVKNFOIVHKNDPDYIYMOGRVADDFVFTLDYNY 420
 DB 361 lhnkapyvnsdtqsyvlnfrgrvtgasvknfivhknbdpylvmqgrvadvdfldyny 420
 QY 421 PLCAVQAFGIGLSSFDKRI 439
 DB 421 plcavqafgiglsfsfksl 439

RESULT 3
 ID AAM36493 standard; Protein; 442 AA.
 XX AAM36493;
 AC AAM36493;
 XX 27-APR-1998 (first entry)
 DT 27-APR-1998 (first entry)
 XX Human TULP3 protein.
 DE Human TULP3 protein.
 XX TULP3; TULP3; sensory neuron; neurosensory defect;
 KM cochlear degeneration; hearing loss; deafness; retinal dystrophy;
 KW retinitis pigmentosa; combined rod cone dystrophy; human;
 KM animal model; transgenic animal; therapy; diagnosis.
 XX Homo sapiens.
 OS Homo sapiens.
 XX MO9738004-A1.
 FN MO9738004-A1.
 XX 16-OCT-1997.
 PD 16-OCT-1997.
 XX 10-APR-1997; 97WO-US05903.
 PE 10-APR-1997; 97WO-US05903.
 XX 17-SEP-1996; 96US-0714991.
 PR 10-APR-1996; 96US-0630592.
 PR 22-AUG-1996; 96US-0701380.
 PR 04-SEP-1996; 96US-0706292.
 XX (JACK-) JACKSON LAB.
 PA (SEQU-) SEQUANA THERAPEUTICS INC.
 XX Nagert J, Nishina P, Noben-Trauth K, North M;
 PI WPI: 1997-512642/47.
 DR N-PSDB: AAT96644.
 XX Mammalian TULP protein - used for detecting pre-disposition to
 PT neuro-sensory defects
 XX Claim 3; Page 59-60; 89pp; English.
 PS This protein comprises human TULP3, a member of the mammalian
 CC TULP family. Its amino acid sequence was deduced from an isolated
 CC cDNA clone (see AAT96644). Expression of TULP2 is restricted to the
 CC retina. TULP3 is a member of the mammalian TULP gene family
 CC associated with various defects in sensory neurons such as cochlear
 CC defects, retinitis pigmentosa and combined rod cone dystrophy.
 CC TULP family polypeptides can be used as immunogens to raise
 CC antibodies that specifically identify TULP expressing cells, in
 CC drug screening assays directed at neurosensory defects, and for
 CC therapeutic purposes.

XX SQ Sequence 442 AA;
 Query Match 82.1%; Score 2109; DB 18; Length 442;
 Best Local Similarity 93.4%; Pred. No. 1,66-203;
 Matches 410; Conservative 3; Mismatches 26; Indels 0; Gaps 0;

QY 1 MEASRCRLSPSGDSYFHEEMMKRQAKLDYORLLERQRKRRLEPPVOPNPEARLRA 60
 DB 1 measrcrlspgsdsyfheemmkrmqakldygrlllekrrkrllepfnvqpnpearlra 60
 QY 61 KPRADEQTPLVNCHTPHSNVLHGIDGPAVLRKPEVHAPESSSVVEDEAENTVDIAS 120
 DB 61 kpradeqtplvnchtphsnvllhgidgpaavlrkpedevhapesvssvveedaentvdtas 120
 QY 121 KPGLOERLQKHDSYVFEDETGISQACLERPNNSQNSTDTSGSATRAQPADN 180
 DB 121 kpglqerlqkhdssyvfeede tgisqaclerpnnsqnsdtgtg1pvl1lppnqllt 180
 QY 181 LLGDDIDLEDFVYSPAPQGVTVRCRIIRDRKGMGRGLFPTYYMYLKEENOKIFLLAARK 240
 DB 181 llgddidledfvysspapgvttvrcrllrdkrgmgrglfshllylvgkeengkflllaark 240
 QY 241 RKSKTANYLISIDPVDLSREGESYVGKLSNLMGTRKFTYDNGICPMKRGILVGAATHR 300
 DB 241 rkksktanyllisidpvdlsregesyvgklrsnlmgtrkftvydngicpmkgqrgilvgaahr 300
 QY 301 QELAAISYETNVLGFGPRKMSVITPGMTLNHKQIPYQPNNDLSLRMONTMENTLVE 360
 DB 301 qelaaissyenvlfgprkmsvllpgmtlnhkqipyqpnmhdsllstrwqntmentlve 360
 QY 361 LHNKAPVNSDTQSYVLNFRGRVTOASVKNFOIVHKNDPDYIYMOGRVADDFVFTLDYNY 420
 DB 361 lhnkapyvnsdtqsyvlnfrgrvtgasvknfivhknbdpylvmqgrvadvdfldyny 420
 QY 421 PLCAVQAFGIGLSSFDKRI 439
 DB 421 plcavqafgiglsfsfksl 439

RESULT 4
 ID AAB26907 standard; Protein; 442 AA.
 XX AAB26907;
 AC AAB26907;
 XX 15-JAN-2001 (first entry)
 DT 15-JAN-2001 (first entry)
 XX Human TULP3 protein.
 DE Human TULP3 protein.
 XX Human; TULP; neurosensory defect; retina; retinal dystrophy;
 KW Chromosome 12p13.2-12p13.3.
 XX Homo sapiens.
 OS Homo sapiens.
 XX US6114502-A.
 PN US6114502-A.
 PD 05-SEP-2000.
 XX 27-FEB-1998; 98US-0032365.
 PE 27-FEB-1998; 98US-0032365.
 XX 22-AUG-1996; 96US-0701380.
 PR 04-SEP-1996; 96US-0706292.
 PR 10-APR-1996; 96US-0630592.
 PR 17-SEP-1996; 96US-0714991.
 PR 30-APR-1997; 97US-0850218.
 PR 01-AUG-1997; 97US-0904699.
 PR 17-SEP-1997; 97US-0932306.
 XX (AXS-) AXS PHARM INC.
 PA (AXS-) AXS PHARM INC.
 XX North M, Nishina P, Noben-Trauth K, Nagert J;

XX WPI: 2000-586483/55.
 DR N-PSDB: AAA94637.
 XX Mammalian proteins expressed in retina and brain, useful for producing
 PT antibodies and for diagnosing neurosensory defects including cochlear
 PT degeneration, peripheral retinal degeneration and cone-rod retinal
 PT dystrophy -

XX Claim 1; Columns 69-72; 61pp; English.

XX The present sequence is human TULP3. The gene encoding this protein is a
 CC member of the neurosensory defect associated gene family, and is
 CC expressed in the retina. The human TULP3 gene maps to chromosome
 CC 12p13.2-12p13.3. The TULP3 gene is useful as an immunogen to raise
 CC antibodies that specifically identify TULP expressing cells and in drug
 CC screening assays directed at neurosensory defects. The present protein
 CC can be used for the treatment of neurosensory degenerative conditions
 CC (retinal dystrophies) e.g. retinitis pigmentosa, combined cone rod
 CC dystrophy, age related macular dystrophy, Stargardt's macular dystrophy,
 CC Best's disease, pigment pattern dystrophies, central alveolar chorioid
 CC dystrophy, dominant drusen, hereditary haemorrhagic macular dystrophy,
 CC North Carolina macular dystrophy, pericentral choroidal dystrophy, adult
 CC foveomacular dystrophy, benign concentric annular macular dystrophy,
 CC central areolar pigment epithelial dystrophy, congenital macular
 CC coloboma, dominantly inherited cystoid macular oedema, familial foveal
 CC retinoschisis, fenestrated sheen macular dystrophy, progressive foveal
 CC dystrophy, slowly progressive macular dystrophy, Sorby's
 CC pseudoinflammatory dystrophy, progressive cone dystrophy, Leber's
 CC congenital amaurosis and Goldman-Favre syndrome.

XX Sequence 442 AA;

Query Match 82.1%; Score 2109; DB 21; Length 442;
 Best Local Similarity 93.4%; Pred. No. 1.6e-203;
 Matches 410; Conservative 3; Mismatches 26; Indels 0; Gaps 0;

QY 1 MEASRCRLSPSGDSVFHEEMMKRQAKLDYORLLERKQRKRLPEFVQNPPEARLRA 60
 Db 1 measrcrlspsgdsvfheemmkrmqakldygrlllekrrkrlpefvmqnppearlrra 60
 QY 61 KPRASDEQTPLVNCHTPHSNVLHGIDGPAVNLKPEVHAPSVSSEVEDAENTVTAS 120
 Db 61 kprasedeqtplvnchtpshnvlhgidgpaavlnkpevhapsvssveedaentvtas 120
 QY 121 KPGLOERLQKHDISESVNFDEETDGISQSACLERPNSASQNSTDTGSSGATTAQPADN 180
 Db 121 kpglgerlqkhdisesvndeetdgisqsaclerpnssasqnsdtgipvlllppnqlt 180
 QY 121 kpglgerlqkhdisesvndeetdgisqsaclerpnssasqnsdtgipvlllppnqlt 180
 QY 181 ILGDIIDLEDFVSPAPGVFRCRIIRDKRGMDGLFPTYYMYLEKENOKIFLLARK 240
 Db 181 ilgdiidledfvspapgvfrcririidrkrgmdglfptyymylekenokifllark 240
 QY 181 flgldiddledfvlvpapgvvtrcrlrdkqmdrglshlllylgkeenkifllark 240
 Db 181 flgldiddledfvlvpapgvvtrcrlrdkqmdrglshlllylgkeenkifllark 240
 QY 241 RKKSNTANVLISIDPVDLSREGESYVGRKLSNLMGTFTVYDRGICPMKRGVGAANTR 300
 Db 241 rkksntanvlisidpvdlsregesyvgrklsnltmgtftvydrigcpmkrgvgaaatr 300
 QY 241 rkksntanvlisidpvdlsregesyvgrklsnltmgtftvydrigcpmkrgvgaaatr 300
 QY 301 QELAISETNVLGFRGKRKMSVITPGMTLNHKOIPYQONNHDSLSRMQNRTEENIVE 360
 Db 301 qelaisetnvlgfrgkrkmsvltpgmtlnhkoipyqonnhdsllsrqnrtemenive 360
 QY 301 qelaisetnvlgfrgkrkmsvltpgmtlnhkoipyqonnhdsllsrqnrtemenive 360
 Db 301 qelaisetnvlgfrgkrkmsvltpgmtlnhkoipyqonnhdsllsrqnrtemenive 360
 QY 361 LNNKAPVWNSDTQSYVLNFRGRTQASVKNFQIVHKNDPDYIVMOFGRVADVFTLLDNY 420
 Db 361 lnnkapvwnsdtqsyvlnfrgrtqasvknfqivhknppdyivmofgrvadvftlldny 420
 QY 421 PLCAVQAFGIGLSSFDKRI 439
 Db 421 plcavqafaisissfskl 439

RESULT 5
 AA81948

ID AA81948 standard; Protein: 460 AA.

XX AA81948;

XX 07-JUL-2000 (first entry)

XX Mouse K-tubby protein sequence.

XX K-tubby; mouse; diagnosis; obesity; blindness; deafness; therapy.

XX Mus sp.

XX JP2000053700-A.

XX 22-FEB-2000.

XX 11-AUG-1998; 98JP-0227063.

XX 11-AUG-1998; 98JP-0227063.

XX (TAKE) TAKEDA CHEM IND LTD.

XX WPI: 2000-353265/31.

XX N-PSDB: AAA07391.

XX A new protein useful as drug for treatment of obesity, blindness and

XX deafness -

XX Claim 1: Fig 1-3; 29pp; Japanese.

XX This sequence represents the mouse K-tubby of the invention.

XX The DNA and protein sequences can be used as a drug for the treatment and

XX prevention of diseases caused by deletion and lowered expression of

XX K-tubby. They can particularly be used to diagnose or treat obesity,

XX blindness, or deafness.

XX Sequence 460 AA;

Query Match 63.0%; Score 1617; DB 21; Length 460;
 Best Local Similarity 68.4%; Pred. No. 7.1e-154;
 Matches 314; Conservative 51; Mismatches 72; Indels 22; Gaps 4;

QY 1 MEASRCRLSPSGDSVFHEEMMKRQAKLDYORLLERKQRKRLPEFVQNPPEARLRA 60
 Db 1 measrcrlspsgdsvfheemmkrmqakldygrlllekrrkrlpefvmqnppearlrra 60
 QY 61 KPRASDEQTPLVNCHTPHSNVLHGIDGPAVNLKPEVHAPSVSSEVEDAENTVTAS 110
 Db 61 kprasedeqtplvnchtpshnvlhgidgpaavlnkpevhapsvssveedaentvtas 110
 QY 111 --DAENTVDTASKRGLOERLQKHDISESVNFDEETDGS-----ISQSACLERPNSAS 160
 Db 111 --daentvdtaskrgloerlqkhdisesvndeetdgs-----isqsaclerpnssas 160
 QY 121 sadgespeetapkpdlqellqkngllssvnydeapdkdeegnlsspsarsesaasq 180
 Db 121 sadgespeetapkpdlqellqkngllssvnydeapdkdeegnlsspsarsesaasq 180
 QY 161 QNSTDTGSSGATTAQPADNLDGDIIDLEDFVSPAPGVFRCRIIRDKRGMDGLFPT 220
 Db 161 qnsdtgtssgataaqpadnldgdiidledfvspapgvfrcririidrkrgmdglfpt 220
 QY 181 kaasetgag--vtaqgdagilevenldfayspaprvtvcktrdkkqmdrllft 238
 Db 181 kaasetgag--vtaqgdagilevenldfayspaprvtvcktrdkkqmdrllft 238
 QY 221 YMYLEKENOKIFLLARRKRSKSTANVLISIDPVDLSREGESYVGRKLSNLMGTFTV 280
 Db 221 ymylekenokifllarrkrskstnvlisidpvdlsregesyvgrklsnltmgtftvy 280
 QY 239 YMYLEKENOKIFLLARRKRSKSTANVLISIDPVDLSREGESYVGRKLSNLMGTFTV 298
 Db 239 ymylekenokifllarrkrskstnvlisidpvdlsregesyvgrklsnltmgtftvy 298
 QY 281 YDRGICPMKRGVGAANTHROELAISETNVLGFRGKRKMSVITPGMTLNHKOIPYQOP 340
 Db 281 ydrigcpmkrgvgaaatrroelaisetnvlgfrgkrkmsvltpgmtlnhkoipyqop 340
 QY 299 ydgvnvpvkaqglvekahrqelaalecyenlvigkprkmsvltpgmmnherlfp 358
 Db 299 ydgvnvpvkaqglvekahrqelaalecyenlvigkprkmsvltpgmmnherlfp 358
 QY 341 NNNHDSLSRMQNRTEENIVELNKAPVWNSDTQSYVLNFRGRTQASVKNFQIVHKNDPD 400
 Db 341 nnhdsllsrqnrtemenivelnnkapvwnsdtqsyvlnfrgrtqasvknfqivhknppd 400
 QY 359 nhesliskwqksmenlllelnkapvnddtqsyvlnfrgrtqasvknfqivhgnndp 418
 Db 359 nhesliskwqksmenlllelnkapvnddtqsyvlnfrgrtqasvknfqivhgnndp 418
 QY 401 YIVMOFGRVADVFTLLDNYPLCAVQAFGIGLSSFDKRI 439
 Db 401 yivmofgrvadvftlldnyplcavqafaisissfskl 439

Db 419 ylvmgfgrvadvfdldynpicalqafajlsfaskl 457

RESULT 6
AAW61159
ID AAW61159 standard; Protein; 390 AA.
XX
AC AAW61159;
XX
XX 28-SEP-1998 (first entry)
XX
DE Human tubby 2 protein.
XX
XX Tubby 2; cloning; screening; substantia nigra cell; probe; diabetes;
KW obesity; atherosclerosis; expressed sequence tag analysis.
XX
OS Homo sapiens.
XX
PN EP844253-A1.
XX
PD 27-MAY-1998.
XX
PF 20-NOV-1997; 97EP-0309380.
XX
PR 06-OCT-1997; 97EP-0307877.
PR 25-NOV-1996; 96GB-0024433.
XX
PA (SMIK) SMITHKLINE BEECHAM PLC.
XX
PI Doe TR, Gloger IS, Hughes S, Terret JA, Testa T;
XX
DR WPI: 1998-274150/25.
DR N-PSDB: AAV36466.
XX
PT New 'Tubby 2' DNA and polypeptide sequences - and cell, antibody and
PT agonist etc.
XX
PS Claim 5; Page 4; 22pp; English.
XX
CC The present sequence represents the human Tubby 2 protein. The
CC polynucleotide sequence encoding this protein was obtained by using
CC standard cloning and screening methods from a cDNA library derived from
CC human substantia nigra cells mRNA, using expressed sequence tag analysis.
CC The Tubby 2 sequence may also be obtained from natural sources such as
CC genomic DNA libraries or may be synthesised by using well known and
CC commercially available techniques. This sequence may be used as a probe
CC for cDNA and genomic DNA, so as to isolate full-length cDNAs and genomic
CC clones encoding Tubby 2 polypeptide. It may also be used to isolate cDNA
CC and genomic clones of other genes that have a high sequence similarity to
CC the Tubby 2 gene. The Tubby 2 gene has been mapped to chromosome 12p13.3.
CC The Tubby 2 nucleotide sequence is useful in the design of protocols for
CC the treatment of diabetes, obesity and atherosclerosis and diagnostic
CC assays for the detection of such conditions.
XX
SQ Sequence 390 AA;

Query Match 57.8%; Score 1484; DB 19; Length 390;
Best Local Similarity 97.7%; Pred. No. 1.4e-140;
Matches 291; Conservative 1; Mismatches 4; Indels 2; Gaps 1;

QY 21 MKMQAKLDYORLLERKRRKRLPEPMPQNPENRARRAKPRASDQTLVNCMPHSN 80
Db 1 mkmgakldyqrrlllekrrkrrkrlpepmpqnppearlrrakprasadqtlvncmphsn 60
QY 81 VILHIGDPAVALKPDENVHAPSVSSVVEEDAEVTVDKAPGLOERLOKHDSSEVNF 140
Db 61 vilhigdpavalkpdevhapsvssvveedaentvdtaskpqlgerlqkhdhsesvnf 120
QY 141 EETDGIQSACLERPNASSONSTDTCTSGSATPAQPADNLDLDDLEDFVYSPAPQV 200
Db 121 eetdgisqaclerpnasssonstdtctsgsataaqpadnllgdldledfyvspapqgv 180

QY 201 TVRCRIIRDKRGMDRGLEPTTYMYLEKEENOKITFLAARKKKKSTANYLISDPVLSR 260
Db 181 tvrcrirdkrgmdrglftpyymylekeengkiflfaarkkkskstantylisdpvlstr 240
QY 261 EGESYVGLRSLNMGTFETVYDRGICPMKGRGLVGAAHTQELAAISYEINVLGFKP 318
Db 241 egesyvgklrslnmgftfydrigicpmkgrglvgaahtrqelaaisysv--aafpgp 296

RESULT 7
AAW36495
ID AAW36495 standard; Protein; 518 AA.
XX
AC AAW36495;
XX
XX 27-APR-1998 (first entry)
XX
DE Human TUB Form 2.
XX
XX TULP; TUB Form 2; sensory neuron; neurosensory defect;
KW cochlear degeneration; hearing loss; deafness; retinal dystrophy;
KW retinitis pigmentosa; combined rod cone dystrophy; obesity; human;
KW animal model; transgenic animal; therapy; diagnosis.
XX
OS Homo sapiens.
XX
PN WO9738004-A1.
XX
PD 16-OCT-1997.
XX
PF 10-APR-1997; 97MO-US05903.
XX
PR 17-SEP-1996; 96US-0714991.
PR 10-APR-1996; 96US-0630592.
PR 22-AUG-1996; 96US-0701380.
PR 04-SEP-1996; 96US-0706292.
XX
PA (JACK-) JACKSON LAB.
PA (SEQU-) SEQUANA THERAPEUTICS INC.
XX
PI Nagert J, Nishina P, Noben-Trauth K, North M;
XX
DR WPI: 1997-512642/47.
DR N-PSDB: AAT96683.
XX
PT Mammalian TULP protein - used for detecting pre-disposition to
PT neuro-sensory defects
XX
PS Claim 3; Page 73-74; 89pp; English.
XX
CC This sequence comprises human TUB Form 2, a member of the TULP
CC protein family that is associated with various defects in
CC sensory neurons such as cochlear defects, retinitis pigmentosa and
CC combined rod-cone dystrophy. 6 cDNA splice variants of TUB have
CC been identified (see AAT96682-87) by 5'RACE and are designated as
CC Form 1 through 6. The encoded proteins have a common C-terminal
CC sequence (AAW36488) and vary in their N-terminal sequences (see
CC AAW36495-97). Forms 1 to 4 have unique N-terminal sequences (see
CC AAW36495-97). Forms 5 and 6
CC vary only in the non-translated cDNA sequences. TUB is expressed
CC at high levels in the brain, eye and testis. Human TULP proteins
CC specifically identify TULP expressing cells, in drug screening
CC assays directed at neurosensory defects, and for therapeutic
CC purposes. The N-terminal domain of TUB has been shown to direct
CC nuclear localisation of the protein.
XX
SQ Sequence 518 AA;

Query Match 46.9%; Score 1204.5; DB 18; Length 518;
Best Local Similarity 49.2%; Pred. No. 3.5e-112;
Matches 250; Conservative 74; Mismatches 111; Indels 73; Gaps 10;

KM animal model; transgenic animal; therapy; diagnosis.
 XX Homo sapiens.
 OS
 XX MO9738004-A1.
 PN
 XX 16-OCT-1997.
 PD
 XX 10-APR-1997; 97WO-US05903.
 PF
 PR 17-SEP-1996; 96US-0714991.
 PR 10-APR-1996; 96US-0630592.
 PR 22-AUG-1996; 96US-0701380.
 PR 04-SEP-1996; 96US-0706292.
 XX
 PA (JACK-) JACKSON LAB.
 PA (SEOU-) SEQUANA THERAPEUTICS INC.
 PI Nagert J, Nishina P, Noben-Trauth K, North M;
 DR WPI: 1997-512642/47.
 DR N-PSDB; AAT96684.
 XX
 XX Mammalian TULP protein - used for detecting pre-disposition to
 PT neuro-sensory defects
 PS Claim 3; Page 75-76; 89pp; English.
 XX
 CC This sequence comprises human TUB Form 3, a member of the TULP
 CC protein family that is associated with various defects in
 CC sensory neurons such as cochlear defects, retinitis pigmentosa and
 CC combined rod-cone dystrophy. 6 cDNA splice variants of TUB have
 CC been identified (see AAT96682-87) by 5'RACE and are designated as
 CC Form 1 through 6. The encoded proteins have a common C-terminal
 CC sequence (AAW36488) and vary in their N-terminal sequences (see
 CC AAW36495-97). Forms 1 to 4 have unique N-terminal sequences (see
 CC vary only in the non-translated cDNA sequences. TUB is expressed
 CC at high levels in the brain, eye and testis. Human TULP proteins
 CC (AAW36489-97) are useful as immunogens to raise antibodies that
 CC specifically identify TULP expressing cells, in drug screening
 CC assays directed at neurosensory defects, and for therapeutic
 CC purposes. The N-terminal domain of TUB has been shown to direct
 CC nuclear localisation of the protein.
 CC
 XX
 XX Sequence 512 AA;

Query Match 46.9%; Score 1203.5; DB 18; Length 512;
 Best Local Similarity 49.7%; Pred. No. 4.3e-112;
 Matches 247; Conservative 73; Mismatches 106; Indels 71; Gaps 9;

QY 11 GDSVFHEEMKMKROAKLDYRLLEKRRKLEPEWQNPPEARLRRAKPRASDEQTP 70
 DB 16 sydsylddegnylqkldrgallegkqkkrqelmvgnadgprsrarqseagap 75
 QY 71 LVNCHTPHSNVIH----- 84
 DB 76 lvesylassgsstsyvgeadslasvq]gatrptapasakrtkaataggggaarkekkg 135
 QY 85 ---GIDGPAVALK-PDEVHAPSVSSVVEED-AENTVDTASKPG-----LOBRLKHD 132
 DB 136 khkgtsgpaalaekseagpvyqlltyvqshadagetaaggsqgqdltaqtmgkq 195
 QY 133 ISESYNPE---ETDGISGACLE---RPNSSASSNSTDGTSGSATAA--OPADNVLG 183
 DB 196 lssmsstfedeedeensssqslnsrltrpsatksrksvreaaapltapeqvd---v 252
 QY 184 DIDDLEDFVYSPAGQVTRCRITRDRKGMGRGLFPTVYMLEKEENOKIFLLAARKKK 243
 DB 253 evqdldeefalrpapqglitkricttdkkgmtrgmpvlyfhlldedgkxvflilagrk 312
 QY 244 SKTANVLTIDPVDLSREGESYVGKLSNLMGTFTVYDRGICPMKGRGLVGAHNT-ROE 302
 ||:||||:| ||| ||:||||:||||:||||:||||:| ||| :| |||

DB 313 sktsnyllsvdptdlsrgdsyigklrlnlmgktftvydngvnpkassstlesqtlrge 372
 QY 303 LAAISTEWNIGEGKPRRMSVLIIPGMLNHKQIPYQPNHDSLSRMQNTMENTVELH 362
 DB 373 laaveyenvlvgfkqprkmsvlyvgmmvhervsirpnehetllarqmntesileld 432
 QY 363 NKAPVWNSDQSYVLNFGGRVTOASVKNFQIVHKNDPPYIYMOGRVADVFYLDVNP 422
 DB 433 nktvpmddtqsyvlnfngvrtgsvknfqlhngndpdylymgfgrvadedvftmdynp 492
 QY 423 CAVQAFGIGLSSFDKRI 439
 DB 493 calgaafalalsfsdskl 509

RESULT 10
 AAB26912
 ID AAB26912 standard; Protein; 512 AA.
 XX
 AC AAB26912;
 XX
 DT 15-JAN-2001 (first entry)
 XX
 DE Human TUB form 3 5'RACE protein product.
 XX
 KW Human; TUB; neurosensory defect; retina; retinal dystrophy.
 XX
 OS Homo sapiens.
 XX
 PN US6114502-A.
 PD 05-SEP-2000.
 XX
 PF 27-FEB-1998; 98US-0032365.
 XX
 PR 22-AUG-1996; 96US-0701380.
 PR 04-SEP-1996; 96US-0706292.
 PR 10-APR-1996; 96US-0630592.
 PR 17-SEP-1996; 96US-0714991.
 PR 30-APR-1997; 97US-0850218.
 PR 01-AUG-1997; 97US-0904659.
 PR 17-SEP-1997; 97US-0932306.
 XX
 PA (AXYS-) AXYS PHARM INC.
 XX
 PI North M, Nishina P, Noben-Trauth K, Nagert J;
 DR WPI: 2000-586483/55.
 DR N-PSDB; AAA94677.
 XX

PT Mammalian proteins expressed in retina and brain, useful for producing
 PT antibodies and for diagnosing neurosensory defects including cochlear
 PT degeneration, peripheral retinal degeneration and cone-rod retinal
 PT dystrophy -
 XX
 PS Claim 1; Columns 99-102; 61pp; English.
 XX
 CC TUB is a member of the neurosensory defect associated gene family, and is
 CC expressed in the retina. There are 6 alternative 5' ends for the TUB
 CC transcript, which lead to different amino acid sequences for the
 CC N-terminus. The present sequence is a protein encoded by one such 5' RACE
 CC product from human TUB. The TUB gene is useful as an immunogen to raise
 CC antibodies that specifically identify TUB expressing cells and in drug
 CC screening assays directed at neurosensory defects. The TUB protein can be
 CC used for the treatment of neurosensory degenerative conditions e.g.
 CC retinal dystrophies.
 CC
 XX
 XX Sequence 512 AA;

Query Match 46.9%; Score 1203.5; DB 21; Length 512;
 Best Local Similarity 49.7%; Pred. No. 4.3e-112;
 Matches 247; Conservative 73; Mismatches 106; Indels 71; Gaps 9;

[illegible]

```

Db      370 vqjetnlgtkgprkmsvlypgmmwhevysiprnehetllarvqnktesielqntk 429
Qy      366 PWNSTQSTVSLNFERGRVYQASVKNFOIVHKNDPDYIVMOFGRVADVFETLLDYNPLCAV 425
Db      430 pvnndtqsyvlnfhgrvctqasvknfgilnhgndpdyimqfgrvaevdfndynpical 489
Qy      426 QAFGIGLSFSDKRI 439
Db      490 qafalastrfskrl 503

RESULT 13
ID      AAW75451
AAW75451 standard. Protein; 506 AA.
XX
AC      AAW75451;
XX
DT      27-APR-1999 (first entry)
XX
DE      Human wild type tub protein.
XX
KW      Human; wild type; tubby; identification; SH2 domain; mammal; obesity;
XX      body weight disorder; cachexia; anorexia.
XX      Homo sapiens.
XX
PN      US5861239-A.
XX
PD      19-JAN-1999.
XX
PF      02-SEP-1997; 97US-0922267.
XX
PR      02-SEP-1997; 97US-0922267.
PR      12-APR-1996; 96US-0631200.
PR      28-MAR-1997; 97US-0829553.
XX
PA      (MILL-) MILLENNIUM PHARM INC.
XX
PI      Kapeller R, Kleya PW, Moore KU;
XX
DR      WPI; 1999-130383/11.
XX      N-PSDB; AAX16702.
XX
PT      Identifying compounds which modulate tub protein activity - by
PT      detecting compounds which alter the interaction of tub protein with
PT      a SH2 containing peptide, used to develop agents for treating e.g.
PT      obesity, cachexia or anorexia
XX
PS      Claim 7; Fig 9A-D; 95pp; English.
XX
CC      This sequence represents the amino acid sequence of the human wild type
CC      "tub" protein. The invention relates to a method for identifying
CC      compounds that modulate tub protein activity, especially its interaction
CC      with proteins containing an SH2 domain. The method can be used for
CC      identifying compounds which modulate tub protein activity for use in the
CC      treatment of mammalian body weight disorders including obesity, cachexia
CC      and anorexia.
XX
SQ      Sequence 506 AA;

Query Match 46.6%; Score 1196.5; DB 20; Length 506;
Best Local Similarity 49.6%; Pred. No. 2.1e-111;
Matches 245; Conservative 73; Mismatches 105; Indels 71; Gaps 9;

Qy      14 SVFHEEMAKMKQALDYQRLLEKROKRRKLEPFMYQPNFEARLRRAKPRASDEQPLVN 73
Db      13 svlddeggnlrgqlrdgralleqkqkkrgeplmwqgnadgprrrrrrgeegaprive 72
Qy      74 CHTFHSNVILH----- 84
Db      73 sylssssstsyvqgeadslasvqlglrtptapasakrlkaaaltaqgqgaarkekkgkhk 132

```

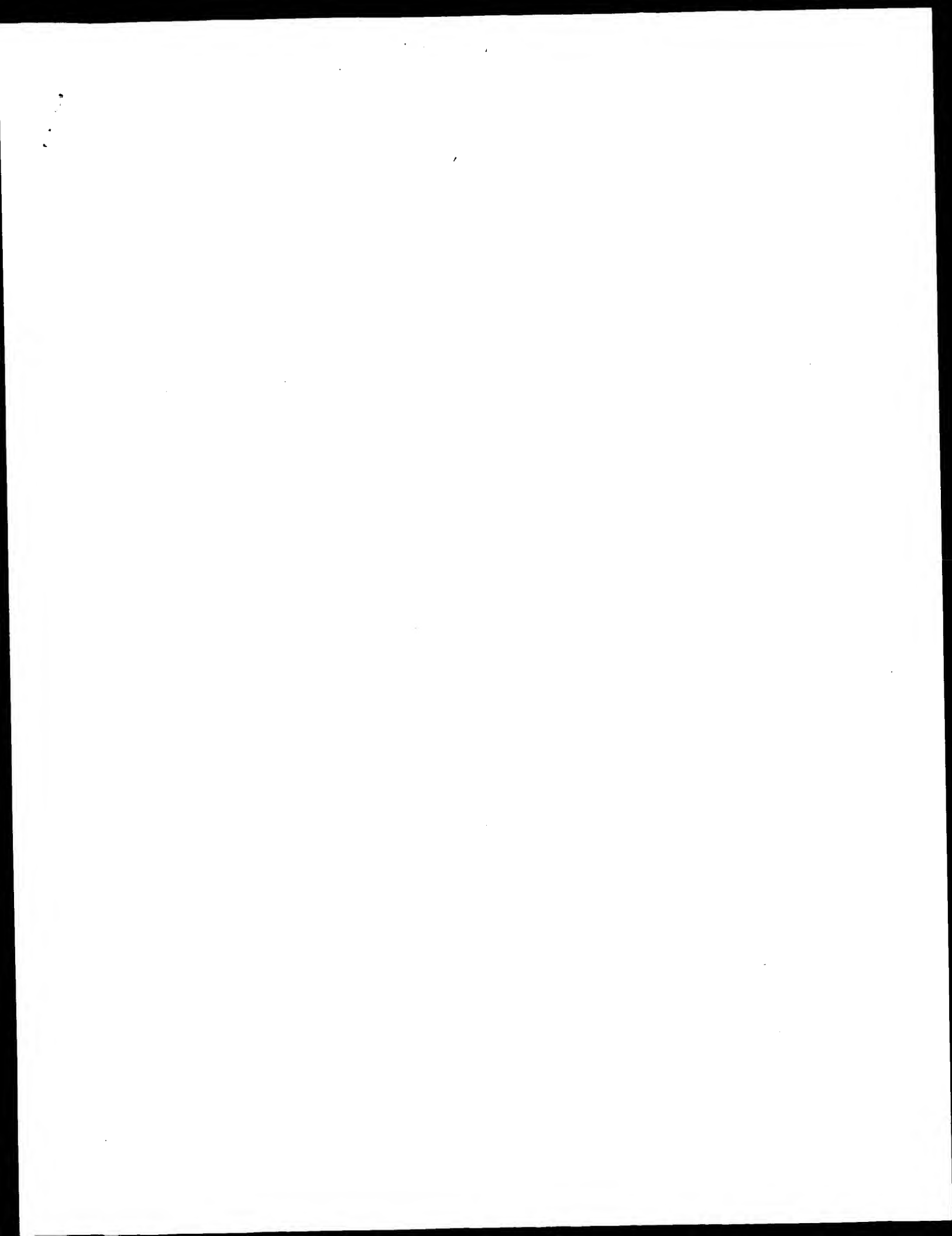

PN W09738004-A1.
 XX 16-OCT-1997.
 PD 10-APR-1997; 97WO-US05903.
 XX 17-SEP-1996; 96US-0714991.
 XX 10-APR-1996; 96US-0630592.
 PR 22-AUG-1996; 96US-0701380.
 PR 04-SEP-1996; 96US-0706292.
 XX (JACK-) JACKSON LAB.
 PA (SEQU-) SEQUANA THERAPEUTICS INC.
 XX Naggert J, Nishina P, Noben-Trauth K, North M;
 PI WPI: 1997-512642/47.
 DR N-PSDB; AAT96640.
 XX Mammalian TULP protein - used for detecting pre-disposition to
 PT neuro-sensory defects
 XX Claim 3; Page 51-52; 89pp; English.
 PS This sequence comprises human TUB Form 1, a member of the TULP
 CC protein family that is associated with various defects in
 CC sensory neurons such as cochlear defects, retinitis pigmentosa and
 CC combined rod-cone dystrophy. 6 cDNA splice variants of TUB have
 CC been identified (see AAT96682-87) by 5'RACE and are designated as
 CC Form 1 through 6. The encoded proteins have a common C-terminal
 CC sequence (AAW36488) and vary in their N-terminal sequences (see
 CC AAW36495-97). Forms 1 to 4 have unique N-terminal forms 5 and 6
 CC vary only in the non-translated cDNA sequences. TUB is expressed
 CC at high levels in the brain, eye and testis. Human TULP proteins
 CC (AAW36489-97) are useful as immunogens to raise antibodies that
 CC specifically identify TULP expressing cells, in drug screening
 CC assays directed at neurosensory defects, and for therapeutic
 CC purposes. The N-terminal domain of TUB has been shown to direct
 CC nuclear localisation of the protein.
 XX
 SQ Sequence 561 AA:

Query Match 46.4%; Score 1192.5; DB 18; Length 561;
 Best Local Similarity 49.5%; Pred. No. 6.4e-11;
 Matches 244; Conservative 73; Mismatches 105; Indels 71; Gaps 9;

OY 15 VFHEEMKMKROAKLDYQRLLEKROKKRLEPFMVOPNPEARLRRAKPRASDEQTPVNC 74
 DB 69 vlddegmrllrqkldtrgalleqkkrgeplmvganadgrrstrarqseegapives 128
 OY 75 HTPHSNVILH-----G 85
 DB 129 ylsstgsstsyvqeadslasvqlatrtptasakrtkaatatggggaarkekgkhg 188
 OY 86 IDGPAAVLK-PDEVHAPSVSSEVVEED-AENTVDTASKPG-----LOERLOKHDSIS 136
 DB 189 tsqpaalaedkseagpvyqlltyvgsdhaqdagetaaggepsgdrlrtmqrkiss 248
 OY 137 VNFDE---ETDGISOSACLE---RPNSSASQNSTDTGTGSATAA--QPADNLIGDIDD 187
 DB 249 msfdeedeedeensssqslnsntpsatslksvreaasaprtapeqvd---vevgd 305
 OY 188 LEDEVYSPADPGVTRCIRIRDKRGMDGLFPYVWYLEKEENOKIFLLAARRKRSKTA 247
 DB 306 leefalrpapgitlcrtrdkkmdrgmrypcyflhldedgkxvflaigrkkskts 365
 OY 248 NYLISIPVDLSRGESYVGLKLSNLMGTFTYDRIQICPMKGRGLVGAHT-ROELAAI 306
 DB 366 nylisvdpdlsrgdsyigklrsnimgtkftlydngvmpqkassstlesgllrgelaav 425
 OY 307 SYETNVLGFGKPRKMSVITIGMTLNHKOIYQPNHDSLSRWOMRTMENLVELHNKAP 366
 |||||

DB 426 cyetnvlfgkprkmsvlypmmnwhevsirprnehetllarwqnkntesilelgnkxp 485
 OY 367 WNSDTSQSYVLENGRGVTRQASVKNQJIVHKNDPXYIYMOGRVADDFYTLIDVYPLCAVO 426
 DB 486 vwnddtgsvylnfngyvtqasvknfqlhngndpdylymqfgrvaedvfmdnyrplcalq 545
 OY 427 AFGIGLSEFDRRI 439
 DB 546 afalaissfdskl 558
 |||||

Search completed: July 16, 2002, 16:11:27
 Job time: 67 sec



Db 371 YETNVLGFGKPRKMSVIVPGMMVHERVCIRPRNEHETLLARWONKTESIELOKTPV 430

QY 368 MNSDPOSYVNLNFRGRTQASVKNFOIVHKNDPDYIVMOFGRVADVEFTLDYNYPLCAVOA 427

Db 431 WNDDDQSYVNLNFRGRTQASVKNFOIHNNDPDYIVMOFGRVADVEFTLDYNYPLCALQA 490

QY 428 FGIQLSSFDKRI 439

Db 491 FAIALSSFDKRL 502

RESULT 2

S42728

phosphodiesterase (clone p4-6) - mouse

C:Species: Mus musculus (house mouse)

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 05-Nov-1999

C:Accession: S42728; I48711

R:Vandutas, V.; Moriguchi, D.J.

Biochim. Biophys. Acta 1217, 203-206, 1994

A:Title: Identification and characterization of the developmentally regulated pattern of

A:Reference number: I48711; MUID:94153998

A:Accession: S42728

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-271 <VAM>

A:Cross-references: EMBL:X69827; NID:q467577; PIDN:CAA49481.1; PID:q467578

A:Note: the sequence of residues 177-215 and the corresponding nucleic acid sequence are

Query Match 30.3%; Score 779; DB 2; Length 271;

Best Local Similarity 55.5%; Pred. No. 9.9e-49;

Matches 141; Conservative 56; Mismatches 57; Indels 0; Gaps 0;

QY 186 DLEDFVYSPAQGYVRCRIIRDKRGMDRGLFPTYYWYLEKEENQKIFLLAARKRRKSK 245

Db 115 EDMEATVLPAPRDHVKQWRIYRNKHGMDKGFPSYVTLBEGDVANHLAARKRRKSK 74

QY 246 TANYLISIDPVDLSREGESYVGLKRSNLMGTKEFTYVDGICPMKGRGLVGAHTRQELAA 305

Db 75 TSNYLSLIDPKMSRNGSNFVGKVSNTVLTGTFITFDNGVNERSYWVPDSARIREELGV 134

QY 306 ISETYNVLCFGKPRKMSVIVPGMMVHERVCIRPRNEHETLLARWONKTESIELOKTPV 365

Db 135 VCETYNVLCFGKPRKMSVIVPGMMVHERVCIRPRNEHETLLARWONKTESIELOKTPV 194

QY 366 PWNNSDPOSYVNLNFRGRTQASVKNFOIVHKNDPDYIVMOFGRVADVEFTLDYNYPLCAV 425

Db 195 FMSDPOSYVNLNFRGRTQASVKNFOIVHKNDPDYIVMOFGRVADVEFTLDYNYPLCAV 254

QY 426 QAFGIGLSSFDKRI 439

Db 255 QAFATLSSFDKRL 268

RESULT 3

T20691

hypothetical protein F10B5.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T20691

R:Sim, M.

submitted to the EMBL Data Library, February 1995

A:Reference number: Z19310

A:Accession: T20691

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-425 <WIL>

A:Cross-references: EMBL:Z48334; PIDN:CAA88311.1; GSPDB:GNO0020; CESP:F10B5.4

A:Experimental source: clone F10B5

C:Genetics:

A:Gene: CESP:F10B5.4

A:Map position: 2

A:Introns: 28/2; 85/2; 151/2; 269/3; 337/1; 407/1

Query Match 30.0%; Score 771.5; DB 2; Length 425;

Best Local Similarity 39.5%; Pred. No. 6.7e-48;

Matches 174; Conservative 72; Mismatches 124; Indels 71; Gaps 10;

QY 34 LLEKRRKKRL-----PFVQPPPEARLARRAKPRASD 66

Db 1 MLEDKQKRRQSGVSRVTTSTAMSNMMDYPTFDSNLFPSISDSSVS-----S 52

QY 67 EQPLVNCHTPHSNVILAGIDGPAAVLKPDEVIAPSSSVSEDEDENYDTASKRGLOE 126

Db 53 MNPPLIPDP-----IAQPRQSNPQPOQVQESLISIGYDPD--DINA 97

QY 127 RLOKHDSSEVNFDEETDGISQACLERPSSASSONSTDTGSGATAQPADNLLGDID 186

Db 98 KLSKVNLTSCVSDDE-----DKRYADSPMTVDVARRISEVLPDVFNK--N 147

QY 187 DLEDFVYSPAQGYVRCRIIRDKRGMDRGLFPTYYWYLEK---EENQKIFLLAARKRK 243

Db 148 NLAFVEDPAAVEHCLYKCSITRQKSGVDMFPTFLHEEFTDKROKIFLLAARKRK 207

QY 244 SKTANYLISIDPVDLSREGESYVGLKRSNLMGTKEFTYVDGICPMKGRGLVGAHTRQEL 303

Db 208 STTANYLISIDPVDLSREGESYVGLKRSNLMGTKEFTYVDGICPMKGRGLVGAHTRQEL 264

QY 304 AAISETYNVLCFGKPRKMSVIVPGMMVHERVCIRPRNEHETLLARWONKTESIELOKTPV 358

Db 265 AAISETYNVLCFGKPRKMSVIVPGMMVHERVCIRPRNEHETLLARWONKTESIELOKTPV 324

QY 359 VELNKAAPWNNSDPOSYVNLNFRGRTQASVKNFOIVHKNDPDYIVMOFGRVADVEFTLDY 418

Db 325 KILSKSPQWMDPOSYVNLNFRGRTQASVKNFOIHQSSPEYIVMOFGRVADVEFTLDY 384

QY 419 NYPLCAVQAFGIGLSSFDKRI 439

Db 385 RYPLSAVQAFGIAMTSFGRL 405

RESULT 4

H84920

probable Tubb family protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: H84920

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bentto, M.I.; Town, C.D.; Fujii, C.Y.

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Yankken, S.E.; Umayam, L.; Tallon,

euse, D.; Nierman, W.C.; White, O.; Eisen, D.A.; Salzberg, S.L.; Frazer, C.M.; Venter

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487

A:Accession: H84920

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-407 <STO>

A:Cross-references: GB:AEO02093; NID:g3738302; PIDN:AAC63644.1; GSPDB:GNO0139

C:Genetics:

A:Gene: At2g47900

A:Map position: 2

Query Match 17.4%; Score 448; DB 2; Length 407;

Best Local Similarity 36.5%; Pred. No. 1.2e-24;

Matches 112; Conservative 45; Mismatches 70; Indels 80; Gaps 10;

QY 195 PAFQGYVRCRIIRDKRGMDRGLFPTYYWYL-----EKENQKIFLLAARKRRKSKTANY 249

Db 116 PGRPSLVQCYIMRNSQ-----TYLYLGINQAASNDGKFLAARFRPTCTDY 169

QY 250 LISIDPVDLSREGESYVGLKRSNLMGTKEFTYVDGICPMKGRGLVGAHTRQEL 303

Db 170 IISLNCDSVSGSNVTYIGKLRNPLGKFTYVDAQPTNPGTQYTRRSRLSLSKOVSPR 229

QY 289 KRGVLGAHNTROELALISYETNVLGFKGRKM-----SVIIPGKTLNKKQIPYOP 339
 Db 230 IFSGNPVVAH-----IYELNVLGSRGRPRMOCYMAIPASAVEPGST-----APRTGT 277
 QY 340 ONNHDSL-----LSRMQNTM-----ENVELHNKAPVNSDPOSYVLF 379
 Db 278 ELVHSMUDSPFSPFSRKSIRAESLPSPPSSNAOKESGLLVKKNAPRHEQLQWCNLF 337
 QY 380 RGHVTOASVKNFOIV--HKND-----DYVMOGRVADVFTLDYNYPLCAVQAQFGL 432
 Db 338 NGRVTVASVKNFOIVAAPENGPAGEHEVNIQLQFGKVGCDVFTMYQYDLSAFQFTICL 397
 QY 433 SSPDKRI 439
 Db 398 SSPDKRI 404

RESULT 5

H96797

hypothetical protein F22K20.1 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: H96797

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Hutzar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marzalla,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719

A:Accession: H96797

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-455 <STO>

A:Cross-references: GB:AE005173; NID:92829918; PIDN:AAC00626.1; GSPDB:GN00141

C:Genetics:

A:Gene: F22K20.1

A:Map position: 1

Query Match

Best Local Similarity 31.8%; Score 427.5; DB 2; Length 455;

Matches 109; Conservative 51; Mismatches 74; Indels 109; Gaps 11;

QY 195 PAPQGVTRCRITIRKRGMDRGLEPTTYMYLEK-----ENOKIFLLAARRKRSKTYNY 249

Db 121 PGRPRATMOCFLTKRDKSNL-----TYHLVLCISPALVLENGK-FLLSAKRIRRTTYEY 173

QY 250 LISIDPVLSREGESYVGKLSNMGTKFTYD-----RGICMKRGVLGAHNT 299

Db 174 VISMADITISRSNTYIGKISNFLTGTFTYDQPAVNSINARVOPV---GLSRFPYS 230

QY 300 RQ-----ELAISYETNVLGFKGRKM-----SVIIPG 327

Db 231 KRVSPKVPSSGYKIAQVSEYELNVLGTRPRRHCHAMNSIPASSLAEGSTVGQDIIYPR 290

QY 328 MTLNH-----KQIPYOPQNNHDS-----LSRMQNTMENVL 361

Db 291 SLIDSFRTISSSSRKITYDYSNDFSSARPSDILGLPSLEQDEYVLEBGERNSPPLY-L 349

QY 362 HNKAPVMSNDPOSYVLANRGRVTOASVKNFOIVHKND----- 398

Db 350 KKKPRRHHQLOQCWCLNFRGRVTVASVKNFOIVAAHNPQOPQOPQOPQOPQOPQOPQ 409

QY 399 --PDYIVMOGRVADVFTLDYNYPLCAVQAQFGLSSPDKRI 439

Db 410 DGPDKIILQFGKVGKDMFTMDFRYPLSAFOAFICLSSPDKRI 452

RESULT 6

E86382

hypothetical protein F4F7.33 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001

C:Accession: E86382

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon

ansen, N.F.; Hughes, B.; Hutzar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marzia

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719

A:Accession: E86382

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-445 <STO>

A:Cross-references: GB:AE005172; NID:911067277; PIDN:AAG28805.1; GSPDB:GN00141

C:Genetics:

A:Map position: 1

Query Match

Best Local Similarity 33.6%; Score 423.5; DB 2; Length 445;

Matches 110; Conservative 49; Mismatches 79; Indels 89; Gaps 11;

QY 195 PAPQGVTRCRITIRKRGMDRGLEPTTYMYLEK-----ENOKIFLLAARRKRSKTYNY 249

Db 123 PGRPRATMOCFLTKRDKSNL-----TYHLVLCISPALVLENGK-FLLSAKRIRRTTYEY 175

QY 250 LISIDPVLSREGESYVGKLSNMGTKFTYD-----MKRGVLGAHNTQ- 301

Db 176 IISMDADNISRSNSYIGKISNFLTGTFTYDQPAVNSINARVOPV---GLSRFPYS 235

QY 302 -----ELAISYETNVLGFKGRKM-----SVIIPG 327

Db 236 SPKVPSSGYKIAQVSEYELNVLGTRPRRHCHAMNSIPASSLAEGSTVGQDIIYPR 290

QY 342 -----NHDSL-----LSRMQNTM-----ENVELHNKAPVNSDPOSYVLF 379

Db 296 LDDSFRTISSSSRKITYDYSNDFSSARPSDILGLPSLEQDEYVLEBGERNSPPLY-L 349

QY 373 QSYVLANRGRVTOASVKNFOIVHKND-----DYVMOGRVADV 412

Db 356 QWCWCLNFRGRVTVASVKNFOIVAAHNPQOPQOPQOPQOPQOPQOPQOPQOPQOPQ 415

QY 413 VFTLDYNYPLCAVQAQFGLSSPDKRI 439

Db 416 MFTMDFRYPLSAFOAFICLSSPDKRI 442

RESULT 7

F96499

hypothetical protein T10P12.9 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: F96499

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon

ansen, N.F.; Hughes, B.; Hutzar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marzia

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719

A:Accession: F96499

E86295
 hypothetical protein T24D18.17 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
 C:Accession: E86295
 R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chlin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maitl, R.; Matzli, R.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719
 A:Accession: E86295
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-397 <STO>
 A:Cross-references: GB:AE005172; NID:g6587813; PIDN:AAFL8504.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1

Query Match 9.4%; Score 242; DB 2; Length 397;
 Best Local Similarity 24.1%; Pred. No. 8.3e-10;
 Matches 92; Conservative 71; Mismatches 159; Indels 60; Gaps 15;

QY 96 DEVHAPSVSSVVEEDAENTVDTASKPGIOERLOKHDISESYNFDEETDG---ISQS-- 149
 DB EVNSPEKYSTVETKRLDLALKSQSMKG-----NSGFPEVYENFKSFGSGRTALKQSSL 88
 QY 150 -ACLEPNASASSONS-TDGTSGATAAPADNLGDIDLEDFYVSPAPQGVTV----- 202
 DB 89 QACMKR-NSEVKSFGMTWTSVDSSEHSSSLAKWEPDSF---AAPSSSTLFRALL 144
 QY 203 -----RCRIIRDK--RGMDRGLEPTYYMYLKEENOKIFLLARKRKSKSTANY 249
 DB 145 CKTLPLDVGRCCTLIYKEQSPGELSHGSVSYSLYTHEGRGRDKRLAVALYHSRNKSKLFR 204
 QY 250 LLSIDPVDLSRGESYVGLKRSNLMGTFTYDRCICPKKGGLVGAHHTROELAISEYE 309
 DB 205 VAQNVKGLICSSDESYSMTANLLGSKYYIWDKGV---RGSVSGKM-VKPLLSVVIET 259
 QY 310 TNLVLFKGP-RKMSYIIPGMLNHKOIYPQONNH-----DSLRSR-ONRTMELVELH 362
 DB 260 PITTWTGSYRMRLLP-----KOQPMOKNNKOVQOASKLPLDMLNNK--EKTQKLC 311
 QY 363 NKAPVNSDTSQSYVINF-----GHVQASVKNFQIVHKNPDYIVMOFGRVADDFETLD 417
 DB 312 SRIPHYNKRISKQHELDPRGRTGLRISQSVKNFQLTLETPTPTIILMGVADKARYVID 371
 QY 418 YNPPLCAVOAFGIGLSSFDKRI 439
 DB 372 FRYPSGYQAFICILASIDSKL 393

RESULT 11
 T15776
 hypothetical protein C34F11.9a - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C:Accession: T15776
 R:Bentley, D.
 submitted to the EMBL Data Library, January 1996
 A:Description: The sequence of C. elegans cosmid C34F11.
 A:Reference number: Z18400
 A:Accession: T15776
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-761 <BND>
 A:Cross-references: EMBL:U46753; NID:g1166623; PID:g1166632; PIDN:AAA85766.1; CESP:C34F11

C:Genetics:
 A:Gene: CESP:C34F11.9a
 A:Introns: 41/1; 73/3; 123/2; 196/1; 228/1; 413/1; 438/2; 479/1; 559/2; 618/1; 693/1

Query Match 5.4%; Score 138; DB 2; Length 761;
 Best Local Similarity 22.8%; Pred. No. 0.067;
 Matches 97; Conservative 62; Mismatches 150; Indels 116; Gaps 21;

QY 35 LEKROKKRLPEFWQNPPEARLRARPRASDEQTPVNCHTPHSN-----VILHGIDGP- 89
 DB 265 LKSRKSRGRVEPEBPAP---FRLKRT-----LTLDLNNHTPAGSWMQGVSNALGGPL 314
 QY 90 -----AAVLKPDVHAPSVSSVVEEDAENTVDTASKPGIOERLOKHDISESYNFDEET 143
 DB 315 KKRUSEGALLPFGTSKA---SDDCNDEGDNTKTRRFLQRRKQSEKSGANGKESG 371
 QY 144 DGISQS-----ACLEPNASASSONS-TDGTSGATAAPADNLGDIDLEDFYVSPAPQ 198
 DB 372 GGSASSSPFGALLIRLSHMAASLTSLISGSKSRSA-----SPSPS 412
 QY 199 -----GVTVRCRIIRDKRGMDRGLEPTYYMYL-----EKENOKIFLLAAR 239
 DB 413 RHRHLEESTIGSESDARVFSDD--DRGPTTDFTSVSRQHEKMAKKKKRNRFRKPSR 470
 QY 240 KRKSKTANYLISID--PVDLSRGESYVGLKRSNLMGTFTYDRCICP---MKG--- 230
 DB 471 ASSFSSTIESMSIDVITVNLMDTVNFIQ--ISIVGQTSNCGDNGIVANIMKGGAVA 527
 QY 291 -RGLVGAHHTROELAISEYE--TN-----VLGFGKGRKMSV-----IIPG 327
 DB 528 LDGRLTAGDMILQVNEFSFENFTNDQAVDLREAVSRGCIKLTVAKSFENGOSCTIP- 586
 QY 328 MTLNHHKQIPIQOPONNHDSLRSWONRTMELVELHNKAPVNSDTSQSYV--NFRGRVQA 386
 DB 587 --RNSREVPAPIDTOAMIOHTNAMRGMPISVE-----GSDVLEWILLDHVEGLRERK 636
 QY 387 SVKNP 391
 DB 637 TARNP 641

RESULT 12
 C84697
 hypothetical protein At2g29510 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: C84697
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bentley, M.I.; Town, C.D.; Fujii, C.Y.M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487
 A:Accession: C84697
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-839 <STO>
 A:Cross-references: GB:AE002093; NID:g3980384; PIDN:AAC95187.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g29510
 A:Map position: 2

Query Match 5.0%; Score 127.5; DB 2; Length 839;
 Best Local Similarity 21.1%; Pred. No. 0.44;
 Matches 92; Conservative 72; Mismatches 161; Indels 111; Gaps 21;

QY 39 ORKKRLPEFWQNPPEARLRARPRAS-----DEQTPVNCHTPHSNVIL-----HGIDG 88
 DB 403 ERARSISPF-----RLISFSIGKSSKSNSTEDAKTP-----PHLSTALLSSRRAGLDN 449
 QY 89 PAAVLKPDVHAPSVSSVVEEDAENTVDTASKPGIOERLOKHDISESYNFDEETDGISQ 148

```

Db 450 PSASSFD-----SSSEFKTSANRGSRSPRLRLDPLIKPKSSHSCRSPESS----- 497
QY 149 SACLERPSSASSONSTDTGSGSATAOPADNLLGDDLEDFYSSAPQ-----GYVRC 204
Db 498 ---LKEAPSSQSSSSFLSRNGKSTVOALFRVTSK-NDOPLETFEAVEKEOSTITAATIRK 553
QY 205 RIRKRGMDGDLPEFTYMYLEKEENQKIFLLAARK-RKKSKTANYLISI---DPVLSR 260
Db 554 QILPEKE--DYGHKTYFTTVOEVQKKAKKMMNSRKVOSQYETINIAQOMKVPSPKPLFL 611
QY 261 EGSEYVGLKRLSMLMGTKFTYVDRGICPMKGRGLVGAHTROELAI-----SYET 310
Db 612 AGE-----RSAENLTLREFVL-----VASESQRTNELAAMVIKIPKLDITTSST 655
QY 311 NVLGRKGRKRSVITPG--MTLNKQIPIYOPQNNHDLISLRWQR-----TME 356
Db 656 TLGDYFAEVNATVYLPSSGVHSLPHKGRF-----SSLQHKWKSDDSCDCGMDTCMLR 708
QY 357 NLVELHNKAPVWNSDTOS--YVLNFRGRVTOASVKNFQIVAKNDPDYIVMOFGHADVDF 414
Db 709 ILTNQHNK-PIIMPSTTSDAKRLFTGGVQE-----NNQNP---YLSFTTYREGVY 755
QY 415 TLIDYNYPLCAVOAFGI 430
Db 756 AVEYNTSLSLQAFSI 771

```

RESULT 13

```

C85093
hypothetical protein AT4g09190 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C:Accession: C85093
R:anonymus, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488
A:Accession: C85093
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-383 <STO>
A:Cross-references: GB:NC_001268; MID:97267562; PIDN:CA878043.1; GSPDB:GN00140
C:Genetics:
A:Gene: AT4G09190
A:Map position: 4

```

Query Match 4.8%; Score 124; DB 2; Length 383;

Best Local Similarity 20.8%; Pred. No. 0.25;

Matches 82; Conservative 49; Mismatches 106; Indels 158; Gaps 19;

```

QY 98 VHAIPSSSVVEDAENYDTASKPGLQERLOKHD-----ISSVNFDEDTGISQACLE 153
Db 53 ITTPPTSSV-----TFSLSRPCVLLIFQKHDKLFFASPVH-----QKTC-- 96
QY 154 RENSASSQSTDTGSGSATAOPADNLLGDDLEDFYSSAPQGVYRCRIIRDKRM 213
Db 97 -EN-----VENFYITIPNNKGKIORCSY----- 118
QY 214 DRGLFTYVYMLEKEENQKIFLLAARKKSKTANYLISIDPVLSRGESEYVGLRSNL 273
Db 119 -HGL-----IYLETSTN-VMFI-----RNPITSFFFLPKLD-SKEGRPLTGLGYDP 163
QY 274 MGKFTYVDRGICPMKGRGLV-----AAH----- 298
Db 164 INKRYV-----LCILKRNKIGILLVGAOESWRLSKGLSHKYTGAKCIDGVYIEG 219
QY 299 -----TROELAISYETNVLGF--KGPRKMSY-----IIPGMTL----- 330
Db 220 SPEDGLRQELAIMSPDLRSKFLILNHPKSSIAATCWSSYEGMLAVSSIASGVSIMILE 279
QY 331 ---NHKQIPIYOPQNNH-DLSLSRWQ-----NRTMENLVELHNKAPVWNSDTOSTYVLNFRGR 382

```

```

Db 280 DADNHQWITKIHPSHREFIKERWKLKGYTRGEFY-----TSYRAYVLNVEGR 329
QY 383 VTQASVKNFQIVHKNNDPDYIVMOFGHADVDFLD 417
Db 330 VLYQWFRILYVDPKRSNRYVHGGIAYDDIRLDD 364

```

RESULT 14

T29327

hypothetical protein M01E11.7a - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T29327

R:Pauley, A.; Gattlung, S.

submitted to the EMBL Data Library, November 1996

A:Description: The sequence of C. elegans cosmid M01E11.

A:Reference number: Z20605

A:Accession: T29327

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1110 <PAU>

A:Cross-references: EMBL:U80450; PIDN:AA837833.1; GSPDB:GN00019; CESP:M01E11.7a

A:Experimental source: strain Bristol N2; clone M01E11

C:Genetics:

A:Gene: CESP:M01E11.7a

A:Map position: 1

A:Insertions: 39/2; 216/2; 266/2; 371/3; 437/2; 501/3; 581/3; 636/2; 659/1; 692/3; 820/1

Query Match 4.5%; Score 115.5; DB 2; Length 1110;

Best Local Similarity 20.8%; Pred. No. 4.9;

Matches 89; Conservative 39; Mismatches 171; Indels 129; Gaps 17;

```

QY 110 EDENYVDTAS-----KPGQERLOKHDISESVNPFDEDTGISQACLE---RPNASSQ 161
Db 107 EKRNSVENVSOEVANTETLPQSTQNDISTPAKMDQVYDQAKSALLEELARAPSAMQH 166
QY 162 NSTDTGSGSATAOPADNLLGDDLEDFYSSAPQGVYRCRIIRDK--RGMRLGLE 218
Db 167 NYWNGEVDNVQVVDQQRAY--ITPSTLQRRPKPPARSGSYRTLDDAYCSDMDLDD 224
QY 219 PTYMYLEKEENQKIFLLAARKKRS-----KTANYLISIDPV-----LSREG 262
Db 225 PEYVL-----NYSNTADLPPEPRQEOHAGTRSVQLPRKKKNFAYVDPDDVLESTKRIG 280
QY 263 ESY-VGLKRSNLMGKFTYVDRGICPMKGRGLVGAHTROELAISYETNVLGFKGPRKM 321
Db 281 SAYSVGVR-----CGGQQQQQEOHNASNDNF----- 307
QY 322 SVIIPGMLNKHQIPIYQ-----PQNNHDLISRWQNRITMENLVELHNKAPVW 368
Db 308 -----SMTLNTPTDYRQHYKRNRCQSVTTPRNHFFSTPREDEA-----DAADTW 353
QY 369 NSDTQSYVLNFRGRVTOASVKNFQIVHKNNDPDY-----VMQGRVADVFTLIDN-YR 421
Db 354 LS-----GKLKRVSKRDIDPDVIRRTQEKMLLEELKDSANNDNQHN 398
QY 422 LCAVOAFGIGLSFPDKRIQTLRMOELCELNROHNSAASLVHRTACQKRWGH-----PWROL 477
Db 399 LPNGHARAGLQND-----PLAEFRREERLANTSPYGEERMRGRMGKPPTP 449
QY 478 POSSLVGP 485
Db 450 PRESSASP 457

```

RESULT 15

T30214

fibrinogen-binding protein - Staphylococcus epidermidis

C:Species: Staphylococcus epidermidis

C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000

C:Accession: T30214

R:Nilsson, M.; Frykberg, L.; Flock, J.I.; Pel, L.; Lindberg, M.; Guss, B.
Infect. Immun. 66, 2666-2673, 1998
A:Title: A Fibrinogen-binding protein of Staphylococcus epidermidis.
A:Reference number: 220781; MUID:98261511
A:Accession: T30214

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1092 <NII>

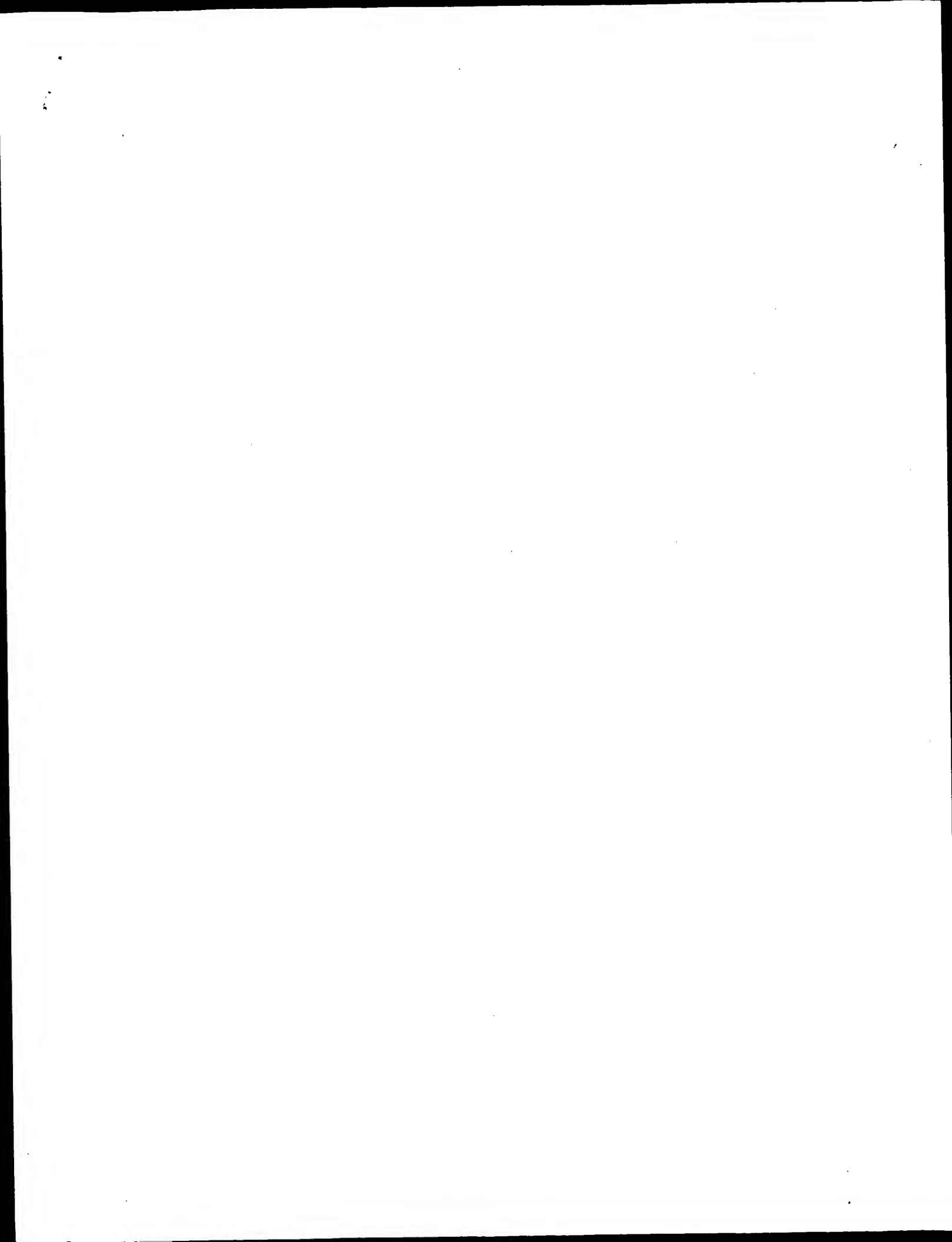
A:Cross-references: EMBL:Y17116; NID:e1296734; PID:e1296735; PIDN:CAA76638.1

Query Match 4.3%; Score 111; DB 2; Length 1092;
Best Local Similarity 19.7%; Pred. No. 10;

Matches 101; Conservative 65; Mismatches 142; Indels 204; Gaps 25;

```
QY 35 LEKRRKKRLPEPMPQNPPEARLRAKPRASDEQPLVNCHPHSNVILHGIDGPAAYLK 94
Db 110 IEKRSDEPRESTNTVDENATFL-----OKTPQDNTH-----LT 143
QY 95 PDEVHAPSVSSVVEEDANVTASKP-----GLQERLQKH-----ISESVNF--D 140
Db 144 EEVEV--KSSSV--ESSNSIDTAQOPSHTTINREESVQTSQDVEDSHVGFANSKIKE 198
QY 141 EETDGISQACLERPNSASSONSTDGTS-----S 171
Db 199 SNTESCKEENTIEQPKV-KEDSTTSQPSGYTNIDEKISNODELNLPINEYENKARPLS 257
QY 172 ATAAQP-----ADNLGDD- 186
Db 258 TTSAGPSIKRVTVNOAABQGSNVNHLIKYTDQSTTEGYDDESGVYKAHDAENLITYTF 317
QY 187 DLEDPYYS-----PAQGVTVRCRI--IDKRG-----MDRGLEPTYM 223
Db 318 EVDQKVSQDNTVIDKNTVPSDLTDSFTIPKIKDNGEIIATGYDNKNKQITTYFTD 377
QY 224 YLEKENQRIFFLAARKRRKSKTANYLISID-----PVDLSREGESYVGL 269
Db 378 YVDKYEENIKAHKLKLSYIDKSKVPPNNNTKLVYEKTKALSSVKNKTTTVEYQRPENRTANL 437
QY 270 RSNLMGTFTFYDRCIGCPKKGGLVGAANTROELAAI-----SYETNVLGFKGPRKMSY 323
Db 438 QS-----METNIDT-----KNHVEQFTIYNPLKYSAKETNNVNISSNGDEGST 480
QY 324 IIPGWTL-----NHKQIPYQPNNDHSLSRNQNTMENTVLELANKAPVWNSDTQSYV 376
Db 481 IIDDSTIIKVKVGDQNL---PDSNRLYDYSEYEDVTINDDYAOLGN-----NNDVN--- 529
QY 377 INFRGRVTQASVKNFQIVHKNDP--DYIVMQ 405
Db 530 INF--GNIDSPYI--IKVISKYDPKDDYTTIQ 558
```

Search completed: July 16, 2002, 16:11:55
Job time: 95 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 16, 2002, 16:11:30 ; Search time 13.43 Seconds
(without alignments)
1415.584 Million cell updates/sec

Title: US-09-782-390-1
Perfect score: 2568
Sequence: 1 MEASRCRLSPSGDSVFHEEM.....HPWRQDPSLVPDLXLM 491

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2109	82.1	442	1	TUL3_HUMAN
2	1617	63.0	460	1	TUL3_MOUSE
3	1197	46.6	505	1	TUB_RAT
4	1196.5	46.6	506	1	TUB_HUMAN
5	1186	46.2	505	1	TUB_MOUSE
6	931	36.3	542	1	TUL1_HUMAN
7	916	35.7	543	1	TULP1_MOUSE
8	899.5	35.0	564	1	TUL2_MOUSE
9	790.5	30.8	520	1	TUL2_MOUSE
10	771.5	30.0	408	1	YQO4_CAEEL
11	193.5	7.5	1547	1	TUSP_MOUSE
12	190.5	7.4	1544	1	TUSP_MOUSE
13	113.5	4.4	1233	1	B3A3_RABIT
14	110	4.3	830	1	YDNL1_SCHPO
15	107.5	4.2	4590	1	FATH_HUMAN
16	107.5	4.2	1232	1	B3A3_HUMAN
17	105.5	4.1	2386	1	FINC_HUMAN
18	104.5	4.1	1739	1	CHD2_HUMAN
19	104	4.0	629	1	PAB2_ARATH
20	104	4.0	751	1	UAS3_DROME
21	103	4.0	710	1	ABBI1_HUMAN
22	102.5	4.0	472	1	WIE_CAEEL
23	102.5	4.0	2035	1	EVPL_MOUSE
24	102	4.0	684	1	SKIL_HUMAN
25	102	4.0	989	1	DLP4_HUMAN
26	102	4.0	1487	1	BLM_DROME
27	100	3.9	722	1	FLID_TREPA
28	100	3.9	1818	1	HMW2_MYCPN
29	100	3.9	2033	1	EVPL_HUMAN
30	100	3.9	2414	1	P300_HUMAN
31	99.5	3.9	568	1	DISC_DROME
32	99.5	3.9	964	1	LON2_MAIZE
33	99.5	3.9	1846	1	MY5B_RAT

34	99	3.9	2150	1	SDC3_CAEEL	P34706 caenorhabdi
35	98.5	3.8	415	1	KAPR_YEAST	P07278 saccharomyc
36	98.5	3.8	586	1	VADL_BPP2	P13520 bacterioph
37	98.5	3.8	719	1	TOP1_MYCGA	O91n65 mycoplasma
38	98.5	3.8	827	1	6P21_YEAST	P40433 saccharomyc
39	98	3.8	321	1	IF2B_SCHPO	P56329 schizosacch
40	98	3.8	708	1	ABY1_MOUSE	O9q4j1 mus musculu
41	98	3.8	1081	1	GALY_YEAST	P19659 saccharomyc
42	97.5	3.8	333	1	CCPA_STRMG	O07329 streptococc
43	97.5	3.8	549	1	MF2_YEAST	P35201 saccharomyc
44	97.5	3.8	1260	1	MY5B_HUMAN	O9u1v0 homo sapien
45	97	3.8	749	1	SWAP_CAEEL	Q10580 caenorhabdi

ALIGNMENTS

RESULT	ID	TUL3_HUMAN	STANDARD	PRT	442 AA.
AC	075386	TUL3_HUMAN			
DT	30-MAY-2000 (Rel. 39, Created)				
DT	30-MAY-2000 (Rel. 39, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Tubby related protein 3 (Tubby-like protein 3).				
GN	TULP3.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE-99047527; PubMed-9828123;				
RA	Nishina P.M., North M.A., Ikeda A., Yan Y., Nagert J.K.;				
RT	"Molecular characterization of a novel tubby gene family member,				
RT	TULP3, in mouse and humans.";				
RL	Genomics 54:215-220(1998).				
CC	-1- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN TESTIS, OVARIES,				
CC	THYROID, AND SPINAL CHORD.				
CC	-1- SIMILARITY: BELONGS TO THE TUB FAMILY.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sdb.ch/announce/				
CC	or send an email to license@isb-sdb.ch).				
CC	-----				
DR	EMBL: AF045583; AAC95431.1; -				
DR	HSSP: P50586; 1C82.				
DR	MTM; 604730; -				
DR	InterPro: IPR000007; Tubby.				
DR	Pfam: PF01167; Tub; 1.				
DR	PROSITE: PS01200; TUB_1; 1.				
DR	PROSITE: PS01201; TUB_2; 1.				
SO	SEQUENCE 442 AA; 49744 MW; 1447469885680283 CRC64;				

Query Match 82.1%; Score 2109; DB 1; Length 442;
Best Local Similarity 93.4%; Pred. No. 4.4e-147;
Matches 410; Conservative 3; Mismatches 26; Indels 0; Gaps 0;

QY	1	MEASRCRLSPSGDSVFHEEMKKRQAKLDYORLLLEKRRKRLPEFMYQPPPEARLRA	60
DB	1	MEASRCRLSPSGDSVFHEEMKKRQAKLDYORLLLEKRRKRLPEFMYQPPPEARLRA	60
QY	61	KPRASDEQRTLVNCHPHSVNVLHAGIDGPAAYAKPPEVAPSSSVVEEDAEVYDTAS	120
DB	61	KPRASDEQRTLVNCHPHSVNVLHAGIDGPAAYAKPPEVAPSSSVVEEDAEVYDTAS	120
QY	121	KPGLQERLQKHCHDISVNEDETDGIGISQACLERPNSASSQNSTDTGTGSAATAQPADN	180
DB	121	KPGLQERLQKHCHDISVNEDETDGIGISQACLERPNSASSQNSTDTGTGSAATAQPADN	180

```

Db 121 KEGLOERLOKHDISSVNFDEETDGISOACLERPNSASONSOTDGTGPVLLLPNOQIT 180
Qy 181 LIGDIDDDLEDFYSPAPOGVTRCIRIDKRGMDGLPPTYMYLEKEENOKIFLLARX 240
Db 181 LIGDIDDDLEDFYSPAPOGVTRCIRIDKRGMDGLPPTYMYLEKEENOKIFLLARX 240
Qy 241 RKRSKTANYLISIDPVDLSREGESYVGLRSLNLMGTFTVYDRGICPKKGRGLVGAHTR 300
Db 241 RKRSKTANYLISIDPVDLSREGESYVGLRSLNLMGTFTVYDRGICPKKGRGLVGAHTR 300
Qy 301 QELAISYETNVLGFKGPKKMSVITPGMTLNKQIPYOPONNHDSLSRMQRTMENIYE 360
Db 301 QELAISYETNVLGFKGPKKMSVITPGMTLNKQIPYOPONNHDSLSRMQRTMENIYE 360
Qy 361 LHNKAPVWMSDTSQYVLNFRGRVTOASYKNFOIVHKNDPDTYVMOFGHADVETLDYNY 420
Db 361 LHNKAPVWMSDTSQYVLNFRGRVTOASYKNFOIVHKNDPDTYVMOFGHADVETLDYNY 420
Qy 421 PLCAVOAFGIGLSSFDKRI 439
Db 421 PLCAVOAFGIGLSSFDKRI 439

RESULT 2
TUB_MOUSE
ID TUB_MOUSE STANDARD: PRT: 460 AA.
AC 088413:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tubby related protein 3 (Tubby-like protein 3).
GN TULP3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN NCBI_TaxID=10090;
RP SEQUENCE FROM N.A.
RA STRAIN-BALB/C;
RX MEDLINE=99047527; PubMed=9828123;
RA Nishina P.M., North M.A., Ikeda A., Yan Y., Naggert J.K.;
RT "Molecular characterization of a novel tubby gene family member,
RT TULP3, in mouse and humans.";
RT Genomics 54:215-220(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN-ICR;
RA Shitayoshi Y., Kawamura A., Nakatsuji N.;
RA Submitted (Aug-1998) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE TUB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC
CC DR EMBL; AF045582; AAC95430.1; -
CC DR EMBL; AB016963; CAB19494.1; -
CC DR HSSP; P50586; 1c82.
CC DR MGD; MGI:1329045; Tulp3.
CC DR InterPro; IPR000007; Tubby.
CC DR Pfam; PF01167; Tdb; 1.
CC DR PROSITE; PS01200; TUB_1; 1.
CC DR PROSITE; PS01201; TUB_2; 1.
CC SEQUENCE 460 AA; 51230 MW; 02EBC62899BB8B3B CRC64;

Query Match 63.0%; Score 1617; DB 1; Length 460;
Best Local Similarity 68.4%; Pred. No. 4,8e-11;
Matches 314; Conservative 51; Mismatches 72; Indels 22; Gaps 4;

```

```

Qy 1 MEASCRFLSPGDSVFHEEMMKRQAKLDYORLLLEKROKRLPEPMYOPNPEARLRA 60
Db 1 MEARCAPAGRGDSAPFDETLRLROLKLDNQRALLEKQKRLLEPLMQPNPEARLRL 60
Qy 61 KPRASDEQTEPLVNCCHPHSNVILHIGDPAVALKPD-----EVAHPSVSSVVEE--- 110
Db 61 KPRSEEHTEPLVDPQMRSDVILHIGDPAALFLKEADLESKPYLSVGSAPAEGETG 120
Qy 111 --DAENTVDPAKRGLOERLOKHDISSVNFDEETG-----ISOACLERPNSASS 160
Db 121 SADSSEPEETAPKPDLEILLQKHGILLSVNYDEEPKEDEGGNLSPSARSESAAQA 180
Qy 161 QNSTDTGSGATPAQPADNLDGIDDEDFYSPAPOGVTRCIRIDKRGMDRLPFT 220
Db 181 KAASETASG--VTAQGDGQOLGEVENLEDPFAKSPARQVTVCKTRDKKGRDLFP 238
Qy 221 YYNTEKEENOKIFLLAARRKRSKTANYLISIDPVDLSREGESYVGLRSLNLMGTFTV 280
Db 239 YYMLEREENKIFLLAGRRKRSKTSNYLVSTPDLRSREGSYIGKLRSLNLMGTFTV 298
Qy 281 YDRGICPKKGRGLVGAHTRQELALISYETNVLGFKGPKKMSVITPGMTLNKQIPYQ 340
Db 299 YDHGVNVPVKAQGLVEKHAHQELALICYETNVLGFKGPKKMSVITPGMNNHRIEFP 358
Qy 341 NNHDSLSRMQRTMENIYELHNKAPVWMSDTSQYVLNFRGRVTOASYKNFOIVHKNDP 400
Db 359 NEHESLSKQKNSKMEWELIELHNKAPVWMSDTSQYVLNFRGRVTOASYKNFOIVH 418
Qy 401 YIVMOFGHADVETLDYNYPLCAVOAFGIGLSSFDKRI 439
Db 419 YIVMOFGHADVETLDYNYPLCALQAFALGSSFDKRI 457

RESULT 3
TUB_MOUSE
ID TUB_MOUSE STANDARD: PRT: 505 AA.
AC 088808: 0921A2;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Tubby protein homolog.
GN TUB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN NCBI_TaxID=10116;
RP SEQUENCE FROM N.A.
RA Miyakita A., Okuno S., Matanabe T.K., Oga K., Tsuji A., Hishigaki H.,
RA Suto T., Nakagawa K., Nakahara Y., Higashi K.;
RT "Molecular cloning of rat tubby gene.";
RT Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE OF 291-404 FROM N.A.
RA STRAIN-WISTAR;
RA Kortschoner N.P., Alvarez-Dolado M., Zenke M.;
RT "Cloning and expression of the rat gene tubby.";
RT Submitted (MAR-1997) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: COULD BE INVOLVED IN THE HYPOTHALAMIC REGULATION OF
CC BODY WEIGHT.
CC -----
CC -1- SIMILARITY: BELONGS TO THE TUB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AB011544; BAA32734.1; -

```



```

371
Db 370 VCEIYVLCFKGPRKMSVIVPGNNVHERVSIRPRNEHETLLARQNKNTESIIELOKNT 429
QY 366 PWNSDTOSYVINFGRVYQASVKNFQIYHKNDPDIYMOGRVADDTLLYNPLCAV 425
Db 430 PWNNDTOSYVINFGRVYQASVKNFQIYHGNDPDIYMOGRVADDTLLYNPLCAL 489
QY 426 QAFGIGLSSFDKRI 439
Db 490 QAFALSSFDKSL 503

RESULT 5
TUB_MOUSE STANDARD: PRT; 505 AA.
ID TUB_MOUSE
AC P50586;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tubby protein.
GN TUB OR RD5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=96200779; PubMed=8612280;
RA Klevn P.W., Fan W., Kovats S.G., Lee J.L., Pulido J.C., Wu Y.,
RA Berlemeyer L.R., Misumi D.J., Holmgren U., Charlat O., Woolf E.A.,
RA Tayber O., Brody T., Shu P., Hawkins F., Kennedy B., Baldwin L.,
RA Ebeling C., Alperin G.D., Deeds J., Lakey N.D., Calpepper J.,
RA Chen H., Gluecksmann-Klus M.A., Carlson G.A., Duyk G.M., Moore K.J.;
RT "Identification and characterization of the mouse obesity gene tubby:
RT a member of a novel gene family.";
RL Cell 85:281-290(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J, AND CD-1; TISSUE=Brain, and Testis;
RX MEDLINE=96195061; PubMed=8606774;
RA Noben-Trauth K., Nagert J.K., North M.A., Nishina P.M.;
RT "A candidate gene for the mouse mutation tubby.";
RL Nature 380:534-538(1996).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 241-505.
RX MEDLINE=20059926; PubMed=10591637;
RA Boggan T.J., Shan W.S., Santagata S., Myers S.C., Shapiro L.;
RT "Implication of tubby proteins as transcription factors by structure-
RT based functional analysis.";
RL Science 286:2119-2125(1999).
CC -1- FUNCTION: COULD BE INVOLVED IN THE HYPOTHALAMIC REGULATION OF
CC BODY WEIGHT.
CC -1- DISEASE: DEFECTS IN TUB ARE THE CAUSE OF MATURITY-ONSET OBESITY,
CC INSULIN RESISTANCE AND SENSORY DEFICITS.
CC -1- SIMILARITY: BELONGS TO THE TUB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U54643; AAB53495.1; -
DR EMBL; U52433; AAC52510.1; -
DR EMBL; U52824; AAC52512.1; ALT_INT.
DR PDB; 1C8Z; 12-DEC-99.
DR MGD; MGI:98868; tub.
DR InterPro; IPR000007; Tubby.
DR Pfam; PF01167; Tub; 1.
DR PROSITE; PS01200; TUB_1; 1.

```

```

DR PROSITE; PS01201; TUB_2; 1.
KW Obesity; Alternative splicing; 3D-structure.
FT DOMAIN 197 204 ASP/GLU-RICH.
FT DOMAIN 206 210 POLY-SER.
FT VARSPPLIC 153 208 MISSING (IN SHORT ISOFORM).
SQ SEQUENCE 505 AA; 55362 MW; 32E160BF5265211 CRC64;

Query Match 46 28; Score 1186; DB 1; Length 505;
Best Local Similarity 49.08; Pred. No. 2e-79;
Matches 241; Conservative 76; Mismatches 107; Indels 68; Gaps 9;

QY 14 SVFHEEMKMKROAKLIDYORLLERKQRKRLEPFMVOPNPEARLRRAKPPASDEQTPLVN 73
Db 13 SYLDDESNLRQKQKLDQRALLBQKQKQKQEPFLMVGANDGRPSRAQSEQAPLVE 72
QY 74 CHTHSNVILH----- 84
Db 73 SYLSSSGSTSYQVEADSIASVOLGATPPAPASAKKSGAASGCGCAPRREKKRKH 132
QY 85 GIDGPAVLK-PDEVHAP---SVSSSVVEEDA-ENIVDTASKPG--IQERLQKHDISE 135
Db 133 GTSGPATLAEDKSRAGPVOILTVGOSDHDKDETRAGGGAQPSQDILKATMQRGKISS 192
QY 136 SVNFEEDTD-----GISQACLERPNSASSONS-TDTGSGSATAPADNLDGIDDL 188
Db 193 SMSFDEDEDEDENSSSSQNSNTRPSATSRSKSIRESAASAPPAPEPPYDI--EVQDL 250
QY 189 EDFYSPAPGVYRCIITIDKRGMDRGLPPTYMYLKEENOKIFLLAARRKRSKTAN 248
Db 251 EEFLRPAPOGITIKRITFDKKGMGRMYPTEVHLHREDGKVPFLGRRKKKSKTSN 310
QY 249 YLISIDPVDSRGESYVGKLRSLNMGTKPTTYVDRCIGPKKGGLVGAHT-ROETLAIS 307
Db 311 YLISVPTDLSRGDSITGLRSLNMGTKPTTYVDNCGVNPOKASSSTLESGLTQRLAAVC 370
QY 308 YETNVLCFKGPRKMSVIVPGNNVHERVSIRPRNEHETLLARQNKNTESIIELOKNT 367
Db 371 YETNVLCFKGPRKMSVIVPGNNVHERVSIRPRNEHETLLARQNKNTESIIELOKNT 430
QY 368 PWNSDTOSYVINFGRVYQASVKNFQIYHKNDPDIYMOGRVADDTLLYNPLCAV 427
Db 431 WNDTOSYVINFGRVYQASVKNFQIYHGNDPDIYMOGRVADDTLLYNPLCAL 490
QY 428 FGIGLSSFDKRI 439
Db 491 FATALSSFDKSL 502

RESULT 6
ID TUB1_HUMAN STANDARD: PRT; 542 AA.
AC 000294; O43536;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tubby related protein 1 (Tubby-like protein 1).
GN TUBP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RX MEDLINE=97250501; PubMed=9096357;
RA North M.A., Nagert J.K., Yan Y., Noben-Trauth K., Nishina P.M.;
RT "Molecular characterization of TUB, TUBP1, and TUBP2, members of the
RT novel tubby gene family and their possible relation to ocular
RT diseases.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:3128-3133(1997).
RN [2]
RP SEQUENCE FROM N.A.

```

RX MEDLINE-98122583; PubMed-9462751;
 RA Baretjee P., Kley P.M., Knowles J.A., Lewis C.A., Ross B.M.,
 Parano E., Kovats S.G., Lee J.J., Penschaszadeh G.K., Ott J.,
 RA Jacobson S.G., Gillingham T.C.;
 RT "TULP1 mutation in two extended Dominican kindreds with autosomal
 RT recessive retinitis pigmentosa";
 RL Nat. Genet. 18:177-179(1998).
 RN [3]
 RP VARIANTS RP14 THR-67 AND THR-489.
 RX MEDLINE-96832025; PubMed-9660588;
 RA Gu S., Lennon A., Li Y., Lorenz B., Fossarello M., North M., Gal A.,
 RA Wright A.;
 RT "Tubby-like protein-1 mutations in autosomal recessive retinitis
 RT pigmentosa";
 RL Lancet 351:1103-1104(1998).
 RN [4]
 RP VARIANTS RP14 PRO-420; LYS-459 AND LEU-491.
 RX MEDLINE-98122582; PubMed-9462750;
 RA Hagstrom S.A., North M.A., Nishina P.M., Berson E.L., Dryja T.P.;
 RT "Recessive mutations in the gene encoding the tubby-like protein TULP1
 RT in patients with retinitis pigmentosa";
 RL Nat. Genet. 18:174-176(1998).
 CC -1- FUNCTION: NOT KNOWN; PROBABLY PLAYS AN ESSENTIAL ROLE IN THE
 CC PHOTORECEPTORS.
 CC -1- TISSUE SPECIFICITY: RETINAL-SPECIFIC.
 CC -1- DISEASE DEFECTS IN TULP1 ARE THE CAUSE OF RETINITIS PIGMENTOSA
 CC TYPE 14 (RP14); AN AUTOSOMAL RECESSIVE FORM OF VISION DEFECT.
 CC -1- SIMILARITY: BELONGS TO THE TUB FAMILY.
 CC -1- DATABASE: NAME-Mutations of the TULP1 gene;
 CC NOTE-Retina International's Scientific Newsletter;
 CC WWW="http://www.retina-international.com/sci_news/tulpmut.htm".
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL; U82468; AAB53700.1; -;
 DR EMBL; AF034923; AAB97966.1; -;
 DR EMBL; AF034919; AAB97966.1; JOINED.
 DR EMBL; AF034920; AAB97966.1; JOINED.
 DR EMBL; AF034921; AAB97966.1; JOINED.
 DR EMBL; AF034922; AAB97966.1; JOINED.
 DR HSSP; P50586; IC8Z.
 DR MIM; 602280; -;
 DR MIM; 600132; -;
 DR InterPro: IPR000007; Tubby.
 DR Pfam: PF01167; Tub; 1.
 DR PROSITE: PS01200; TUB_1; 1.
 DR PROSITE: PS01201; TUB_2; 1.
 KM Vision; Retinitis pigmentosa; Disease mutation.
 FT DOMAIN 115 131 POLY-GLU.
 FT DOMAIN 248 254 R -> T (IN RP14).
 FT VARIANT 67 67 /FTID=VAR_008274.
 FT VARIANT 246 246 R -> V (IN RP14).
 FT VARIANT 259 259 /FTID=VAR_008275.
 FT VARIANT 261 261 I -> T (IN RP14).
 FT VARIANT 261 261 N -> T (IN RP14).
 FT VARIANT 378 378 /FTID=VAR_008277.
 FT VARIANT 378 378 R -> H (IN RP14).
 FT VARIANT 420 420 /FTID=VAR_008278.
 FT VARIANT 420 420 R -> P (IN RP14).
 FT VARIANT 454 454 /FTID=VAR_007941.
 FT VARIANT 454 454 T -> M (IN RP14).
 FT VARIANT 459 459 /FTID=VAR_008279.
 FT VARIANT 459 459 I -> K (IN RP14).
 FT VARIANT 489 489 /FTID=VAR_007942.
 FT VARIANT 489 489 K -> T (IN RP14).

FT FT /FTID=VAR_008280.
 FT VARIANT 491 491 F -> L (IN RP14).
 FT VARIANT 496 496 /FTID=VAR_007943.
 FT VARIANT 496 496 A -> T (IN RP14).
 FT VARIANT 496 496 /FTID=VAR_008281.
 SQ SEQUENCE 542 AA; 60650 MW; EC94308E0FF9B3E9 CRC64;
 Query Match 36.3%; Score 931; DB 1; Length 542;
 Best Local Similarity 45.2%; Pred. No. 1e-60;
 Matches 194; Conservative 76; Mismatches 117; Indels 42; Gaps 9;
 QY 36 EKRRKRLPEFVQNPPEARLRRA--KPRASDEQPTLVNCHTPSHNVLHIDPAAV 92
 Db 128 EEEAESEKKEKILLPPKPLREKSSADLKRRRAKAGSPREDLSP-----DPPPK 178
 QY 93 LKPEVHAPS-----VSSVVEEDAEINTVDTASK-PGL-----QERLQK 130
 Db 179 LRVNKKZAPAGEGKMKRTRKSGSEADKDPSSPASARKSPAMFLVGGSPDKALKK 238
 QY 131 HDISESVNPDEPDGISQACLERPNSASSQNSDGTSGSATAAPADNLGDDIDLED 190
 Db 239 KGPFGARKKEEER--EEAATYIKNSQKGAKGKAKKAKKERAPSPV--EVDPEPE 293
 QY 191 FVSPAPQGYVRCRIIRDRKMDGILPYYMYLKEENQKITLLAARKRKSTANYL 250
 Db 294 FVLKPAQGRVTRCRLRDRKMDGMYPSYFLHIDE--KVFLLAGRRKRKSTANYL 351
 QY 251 ISIDPVLRSRGESYVKGKLSNLMGTFTYDNGICPMKRGVLGAHTQELAAISYET 310
 Db 352 ISIDPVLRSRGENFIKRLSNLNGRFTYVDNGQNPGRYS--TNAASLQELAAIYET 410
 QY 311 NVLGFGRPKKSVITPMTLNHKOIPYQPNNDLSLRQNPMTVELHNKAPVNS 370
 Db 411 NVLGFGRPKKSVITPMTLNHKOIPYQPNNDLSLRQNPMTVELHNKAPVNS 470
 QY 371 DTQSVLNFGRKYQASVKNFQIVHKNDPDYIYQGFRAVDVFTLDYNYPLCAVAFGI 430
 Db 471 DSGSYTLNFGGRVYQASVKNFQIVHADDPYIYQGFRAVDVFTLDYNYPLCAVAFGI 530
 QY 431 GLSFPKRI 439
 Db 531 ALSFPDKRL 539
 RESULT 7
 TULP1_MOUSE STANDARD; PRT; 543 AA.
 AC Q92273;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Tubby related protein 1 (Tubby-like protein 1).
 GN TULP1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN-BALB/C;
 RC Naggett J.K., Nishina P.M., Fitch D., McGinnis N., Basson M., Yan G.,
 RA Cardon L., Shiva N., Duyao M., Ikeda A., McGinnis A., North M.A.;
 RT "Molecular evolution of the tubby gene family";
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN-BALB/C;
 RC Ikeda S., Sorokina I., Naggett J.K., North M.A., Nishina P.M.;
 RA "Apoptotic photoreceptor cell death in tubby mice and the localization
 RT of tubby gene family members in the retina";
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: NOT KNOWN; PROBABLY PLAYS AN ESSENTIAL ROLE IN THE


```

Db      327  RIVRNKHGMQGMPPSYVYILEGDEGVAAHFLLAGRKRRKSTSYNLLISLDPKDRMSNG 386
QY      265  YVGLRNSLMQTKTYYVDROGICPMKKGKGLVCAATHRQELAAISYETNVLGKGRKKSVI 324
Db      387  FVGVRSNVSLGTKEFTIEDNGVNPERSYVWPDSARIREELGVACYETNVLGFRKKMTVI 446
QY      325  IPGTLLNKKQLEPYRPNNDHSLSRMQRTNENVELHNKAPVNNSPDQSYVLNFRGRVY 384
Db      447  LFGHDSRKQRRKRVQPNQDSILSRVQKGAGHGILLQNKRPMSMSDESGAVALNFRGRVT 506
QY      385  QASVKNFQIVYHKNDDYIVMGFRVADVFTLDYVNLPCAVQAFGIGISFDKRI 439
Db      507  RASVKNFQIVHAPDEPDHLYLQFGRVAPITFTMDRFRPLCPLOAFALCISFDGRL 561

RESULT 9
TUL2_HUMAN STANDARD; PRT; 520 AA.
AC 000295;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE Tubby related protein 2 (Tubby-like protein 2).
GN TUP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RC TISSUE=Brain;
RC MEDLINE=97250501; PubMed=9096357;
RX North M.A., Naggert J.K., Yan Y., Noben-Trauth K., Nishina P.M.;
RA "Molecular characterization of TUB, TUP1, and TUP2, members of the
RT novel tubby gene family and their possible relation to ocular
RL diseases."
RL Proc. Natl. Acad. Sci. U.S.A. 94:3128-3133(1997).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN RETINA AND TESTIS.
CC -1- SIMILARITY: BELONGS TO THE TUB FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U82469; AAB53701.1; -.
DR HSSP; P50586; IC8Z.
DR MIM; 602309; -.
DR InterPro; IPR000007; Tubby.
DR Pfam; PF01167; Tub; 1.
DR PROSITE; PS01200; TUB_1; 1.
DR PROSITE; PS01201; TUB_2; FALSE_NEG.
SQ SEQUENCE 520 AA; 58640 MW; 9C76363A537EB6EC CRC64;

Query Match 30.8%; Score 790.5; DB 1; Length 520;
Best Local Similarity 35.4%; Pred. No. 1.9e-50;
Matches 188; Conservative 78; Mismatches 152; Indels 113; Gaps 12;

QY 8 LSPSGSYVNE-----EMMKRQAKLDYQALLLEKRRKRRLPEPVQVPPKAR----- 56
Db 1 MSODNDTLMRDITGLIETLAAIRLQKLEQRRLEKRRKQROBELIMVQANDASPWMLRSC 60
QY 57 LRR-----AKPRASDECP-----LVNC-----HTPSNVITLH- 84
Db 61 LPEERLLGDRGLGNPLRLKKYSEAHLPSSIHSLGIVSGCGDRGRGLPIRTETAVFRN 120
QY 85 -GIDGPAVALKPD---EYVADSVSSSV-----EED 111

```

```

Db 121 LGLOSPFLSWLPDMSDALELEVSVENGSVSPPEPKQSPRIKRGQMAHOPRTAAEGESD 180
Qy 112 AENTVDTASKP--GLQERLQKHDISESYNDEFTDGISQSCLEPRPSASSONSTDTGTS 169
Db 181 SQQMDGDAHRSPNMGPNPQMGDDCYENLAFQKEED-----LEKKREASE---STGTN 229
Qy 170 GSATAAQAOPADNL-----GDIDLEDPFYSPAQCVYTRCR 205
Db 230 SSAAHNELSKALKEGEGTSDSDHMRHEASLAIRSPCGLEEDMAMAYULRALPBTMMQCY 289
Qy 206 IIRDKRGMDRGLFPTYYWYLEKEENQKIFLLAAKRRKSKTANYLISIDPVDSREESY 265
Db 290 LTRDHGVDKGLFPLYLYLYLETSLSQRFLLAGKRRRSKTSNLYLISIDPILLSRDDNF 349
Qy 266 VGXLRSLNMGKFLPYVDGICPMKGRGLVGAATHROELAAISYTYANLGFPGPKMAYII 325
Db 350 VGRKASNFVSTKFTFLFDGVPNDREHLTRNTARLRQELGAVCYPPNVLGIGYPRKMTYIL 409
Qy 326 PGMFLNHKQIIPYOPDNHDSLLSRWONTMENTVELHNKAPVMSDTSQSYVLFNRGATQ 385
Db 410 PGTNSQNRINWQPLNEGESSILSRQDRDQKGLLILHNKTPSKMCKENGVTYLNHGRVTR 469
Qy 366 ASYKRFQYVHKNDPRYIYMGFRGADVFTLIDYNNPLCAVAFGIGSSFD 436
Db 470 ASYKRFQYVHKNDPRYIYMGFRGADVFTLIDYNNPLCAVAFGIGSSFD 520

RESULT 10
YQ04_CAEEL
ID YQ04_CAEEL STANDARD: PRT: 408 AA.
AC 009306;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 46.2 kDa protein F10B5.4 in chromosome III.
GN F10B5.4
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Stimms R.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE TUB FAMILY.
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: 248334; CAB61010.1; -.
DR HSSP: P50586; IC8Z
DR Wormpep: F10B5.4; CE01546.
DR InterPro: IPR000007; TubbY.
DR Pfam: PF01167; Tubb_1.
DR PROSITE: PS01200; TUB_1; 1.
DR PROSITE: PS01201; TUB_2; 1.
DR KW Hypothetical protein.
SQ SEQUENCE 408 AA; 46192 MW; 84F50C6F90618181 CRC64;

Query Match 30.0%; Score 771.5; DB 1; Length 408;
Best Local Similarity 39.5%; Pred. No. 3.2e-49;
Matches 174; Conservative 72; Mismatches 124; Indels 71; Gaps 10;

Qy 34 LLEKRRKKRRL-----PFMWQNPPEARLRRAPRASD 66
1 MEDKKQKQRHQSAGSVTRTTSTAMSMNSKKDYPTEDNSLPFSISIDNSSVSY-----S 52

```


DR PROSITE: PS00678: WD_REPEATS_1; FALSE_NEG.
 DR PROSITE: PS50082: WD_REPEATS_2; 1.
 DR PROSITE: PS50294: WD_REPEATS_REGION; 1.
 KW Repeat; WD repeat.
 FT REPEAT 80 119 WD 1.
 FT REPEAT 123 162 WD 2.
 FT REPEAT 165 204 WD 3.
 FT DOMAIN 368 412 SOCS.
 FT DOMAIN 1467 1544 TUB.
 SQ SEQUENCE 1544 AA; 169173 MW; 3520875841BA3BF3 CRC64;

Query Match
 Best Local Similarity 29.5%; Pred. No. 7.7e-06;
 Matches 52; Conservative 26; Mismatches 67; Indels 31; Gaps 3;

OY 293 LVGAATHROELAAISYENVLGFKGPRK-----MSVITPGTILNHKOIPYOPQNNHDSL 346
 DB 1372 LISSPHLGREKKKVKSQDKLTKLTNFEQDSESEPELFISGDELMQSGSRKGM 1431
 OY 347 LSRQNRRTMENVE-----LNKKAPVWNSDQSYVLNFRGRVT 384
 DB 1433 KSKRSPRAGLEAKCRASERKEDGLSGQGVYVWANKQPLMNEATQYQDLDFGGRVT 1491
 OY 385 QASVKNFQIVHKNDDPYVMQFGRVADVFETLDYVNPICAVQAFGIGLSFDFRQ 440
 DB 1492 QESAKNFQIELEGRO---VMQFGRIDGSAYIILDFQYPSAVQAFVALANTYQRLK 1544

RESULT 13
 B3A3_RABIT STANDARD; PRT; 1233 AA.
 ID 018917:
 AC 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Anion exchange protein 3 (Neuronal band 3-like protein) (Anion
 DE exchanger 3 brain isoform).
 GN SLCA43 OR AE3.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Abuladze N., Pushkin A., Kurtz I.;
 RL Submitted (OCT-1997) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: PLASMA MEMBRANE ANION EXCHANGE PROTEIN.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: NEURONAL.
 CC -1- SIMILARITY: BELONGS TO THE ANION EXCHANGER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AF031650; AAB86859.1; -
 DR HSSP: P02730; 1BTR.
 DR InterPro: IPR001717; Anion_exchanger.
 DR InterPro: IPR003020; HCO3_cotransp.
 DR Pfam: PF00955; HCO3_cotransp. 1.
 DR PRINTS: PR01231; HCO3TRNSPORT.
 DR PROSITE: PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE: PS00220; ANION_EXCHANGER_2; 1.
 KW Transmembrane; Glycoprotein; Anion exchange; Lipoprotein; Palmitate.
 FT DOMAIN 1 709 CTIOPLASMIC.
 FT DOMAIN 710 1233 MEMBRANE (ANION EXCHANGE).
 FT TRANSMEM 710 732 POTENTIAL.

FT TRANSMEM 738 775 POTENTIAL.
 FT TRANSMEM 795 817 POTENTIAL.
 FT TRANSMEM 827 848 POTENTIAL.
 FT TRANSMEM 894 911 POTENTIAL.
 FT TRANSMEM 912 926 POTENTIAL.
 FT TRANSMEM 927 947 POTENTIAL.
 FT TRANSMEM 981 1003 POTENTIAL.
 FT TRANSMEM 1029 1050 POTENTIAL.
 FT TRANSMEM 1084 1129 POTENTIAL.
 FT TRANSMEM 1156 1192 POTENTIAL.
 FT CARBOHYD 874 874 N-LINKED (GLCNAC...) (POTENTIAL).
 FT LIPID 1166 1166 PALMITATE (BT SIMILARITY).
 FT DOMAIN 74 83 HIS-RICH.
 FT DOMAIN 136 150 POLY-GLU.
 FT DOMAIN 203 208 POLY-SER.
 FT DOMAIN 309 312 POLY-LYS.
 FT DOMAIN 443 446 POLY-SER.
 FT DOMAIN 1185 1188 POLY-LEU.
 SQ SEQUENCE 1233 AA; 135759 MW; AE486423E9818583 CRC64;

Query Match
 Best Local Similarity 21.3%; Pred. No. 2.4; Length 1233;
 Matches 83; Conservative 54; Mismatches 152; Indels 101; Gaps 19;

OY 36 EKRRKRLEPFMYQNPPEARLRRAK--PRASDEQTPLVNCHTPHSNVILHGIDGPAVALK 94
 DB 141 EEEBEGSEFENVEPPPSGQKAKFSISDED-----DSP-----GLGKAFAFK 187
 OY 95 P-----DEVHAPSVSSSVVEEDAENTVDYDASKPGLQERLQKHIDSEVNDDEETDGI 146
 DB 188 PLPSVGRSDKSPORSVSSSPRARAPRVAGERSRPW-----SPASVD-----L 232
 OY 147 SOSACLERPNSASSONSDTGTSGSATAAPAD---NLG--DIDDEFFVSPAPQV 200
 DB 233 KERIC--PGSA-----LGNPGPEQVPTDAEQLMGLSADLDKMKHRELDN--- 279
 OY 201 TVRCRIIRDK-----RGMDRGLFPYYWLEKE-----ENOKITLLAARRKK 243
 DB 280 GVRRLVKEPVRVQGRSGRLPTLRKKKKQOPDRPHEVVEINELMLDRSGPWH 339
 OY 244 SKTANYLISIDPV--DLRSBEGSYGKI--RSNLMGTRKFTYDNGICPMKGRGLVGAHT 299
 DB 340 RETARWIKFEEDVEETETRMGKPHVASLSFSLLELRTTAHGAALLDLEQTLPLGIAHL 399
 OY 300 RQELAAISYENVNLGFKGPRKMSVITPGMTLNH-----KOIYPOPN-----N 342
 DB 400 VETWIVSDQIR-----PEDRASVLTLLKHKHPNDKDCGSFPNRPSSSVNSVLGN 453
 OY 343 HDSLLSRQNRRTMENIV-ELHNKAPVWNSD 371
 DB 454 HHATPSHGPDGAVPTMADDLGEPAPLWPHD 483

RESULT 14
 YDNL_SCHPO STANDARD; PRT; 830 AA.
 ID YDNL_SCHPO
 AC P87293:
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 91.7 kDa protein Cl6A10.01 in chromosome I.
 GN SPAC16A10.01.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-977;
 RA Murphy L., Harris D., Wood V., Barrell B.G., Rajandream M.A.;
 RL Submitted (JUN-1997) to the EMBL/Genbank/DBJ databases.

DT	01-MAR-2002 (Rel. 41, Last annotation update)
DE	Cadherin-related tumor suppressor homolog precursor (Fat protein
DE	homolog).
GN	FAT.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX	NCBI_TaxId=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Lymphocytes;
RX	MEDLINE=96163873; PubMed=8586420;
RA	Dunne J., Hardy A.M., Poulson R., Jones T.A., Sheer D., Chin W.G.,
RA	Da S.M., Zhao Q., Beverley P.C.L., Owen M.J.;
RT	"Molecular cloning and tissue expression of FAT, the human homologue
RT	of the Drosophila fat gene that is located on chromosome 4q34-q35 and
RT	encodes a putative adhesion molecule.";
RL	Genomics 30:207-223(1995).
CC	-!- FUNCTION: COULD FUNCTION AS A CELL-ADHESION PROTEIN.
CC	-!- SUBCELLULAR LOCATION: Type I membrane protein (by similarity).
CC	-!- TISSUE SPECIFICITY: EXPRESSED IN MANY EPITHELIAL AND SOME
CC	ENDOTHELIAL AND SMOOTH MUSCLE CELLS.
CC	-!- SIMILARITY: CONTAINS 34 CADHERIN DOMAINS.
CC	-!- SIMILARITY: CONTAINS 5 EGF-LIKE DOMAINS.
CC	-!- SIMILARITY: CONTAINS 1 LAMININ G-LIKE DOMAIN.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; X87241; CAA60685.1; -.
DR	HSSP; P09803; IEDH.
DR	MIM; 600976; -.
DR	InterPro: IPR000152; Asx_hydroxyl.
DR	InterPro: IPR002126; Cadherin.
DR	InterPro: IPR000561; EGF-like.
DR	InterPro: IPR000742; EGF_2.
DR	InterPro: IPR001881; EGF_CA.
DR	InterPro: IPR001791; Laminin_G.
DR	Pfam; PF00028; EGF; 3.
DR	Pfam; PF00008; EGF; 5.
DR	Pfam; PF00054; laminin_G; 1.
DR	PRINTS; PR00205; CADHERIN.
DR	SMART: SM00112; CA; 32.
DR	SMART: SM00179; EGF_CA; 1.
DR	SMART: SM00001; EGF-like; 3.
DR	SMART: SM00282; LamG; 1.
DR	PROSITE; PS00010; ASX_HYDROXYL; 1.
DR	PROSITE; PS00232; CADHERIN_1; 16.
DR	PROSITE; PS50268; CADHERIN_2; 33.
DR	PROSITE; PS00022; EGF_1; 4.
DR	PROSITE; PS01186; EGF_2; 1.
DR	PROSITE; PS01187; EGF_CA; 1.
DR	PROSITE; PS50025; LAM_G_DOMAIN; 1.
RW	Cell adhesion; Signal; Glycoprotein; Transmembrane; Calcium-binding;
RW	Repeat; EGF-like domain.
KW	SIGNAL
FT	1 21
FT	CHAIN
FT	22 4590
FT	POTENTIAL.
FT	CADHERIN-RELATED TUMOR SUPPRESSOR
FT	HOMOLOG.
FT	EXTRACELLULAR (POTENTIAL).
FT	POTENTIAL.
FT	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN
FT	4205 4590
FT	DOMAIN
FT	22 149
FT	DOMAIN
FT	150 256
FT	DOMAIN
FT	257 361
FT	DOMAIN
FT	362 463
FT	DOMAIN
FT	464 569
FT	DOMAIN
FT	570 716
FT	DOMAIN
FT	717 822
FT	DOMAIN


```

FT DOMAIN 823 927 CADHERIN 8.
FT DOMAIN 928 1034 CADHERIN 9.
FT DOMAIN 1035 1138 CADHERIN 10.
FT DOMAIN 1139 1245 CADHERIN 11.
FT DOMAIN 1246 1345 CADHERIN 12.
FT DOMAIN 1346 1456 CADHERIN 13.
FT DOMAIN 1457 1562 CADHERIN 14.
FT DOMAIN 1563 1670 CADHERIN 15.
FT DOMAIN 1671 1769 CADHERIN 16.
FT DOMAIN 1770 1882 CADHERIN 17.
FT DOMAIN 1883 1982 CADHERIN 18.
FT DOMAIN 1983 2084 CADHERIN 19.
FT DOMAIN 2085 2185 CADHERIN 20.
FT DOMAIN 2186 2286 CADHERIN 21.
FT DOMAIN 2287 2393 CADHERIN 22.
FT DOMAIN 2394 2495 CADHERIN 23.
FT DOMAIN 2496 2599 CADHERIN 24.
FT DOMAIN 2600 2705 CADHERIN 25.
FT DOMAIN 2706 2811 CADHERIN 26.
FT DOMAIN 2812 2920 CADHERIN 27.
FT DOMAIN 2921 3024 CADHERIN 28.
FT DOMAIN 3025 3127 CADHERIN 29.
FT DOMAIN 3128 3232 CADHERIN 30.
FT DOMAIN 3233 3337 CADHERIN 31.
FT DOMAIN 3338 3442 CADHERIN 32.
FT DOMAIN 3443 3546 CADHERIN 33.
FT DOMAIN 3547 3649 CADHERIN 34.
FT DOMAIN 3650 3790 EGF-LIKE 1.
FT DOMAIN 3791 3827 EGF-LIKE 2.
FT DOMAIN 3831 4011 EGF-LIKE 3.
FT DOMAIN 4012 4051 EGF-LIKE 4.
FT DOMAIN 4052 4089 EGF-LIKE 5.
FT DOMAIN 4127 4163 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 40 40 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 660 660 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 740 740 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 791 791 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 998 998 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1426 1426 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1551 1551 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1751 1751 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1867 1867 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1905 1905 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1943 1943 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1994 1994 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2328 2328 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2467 2467 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3326 3326 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3424 3424 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3446 3446 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3615 3615 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3642 3642 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3718 3718 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4154 4154 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 4590 AA: 506273 MW: 044830CCDD00E60A7 CRC64;

```

Query Match 4.28; Score 108.5; DB 1; Length 4590;
 Best Local Similarity 18.98; Pred. No. 38;
 Matches 102; Conservative 64; Mismatches 208; Indels 167; Gaps 22;

```

QY 62 PRASDEQPTLVNCHTSPHSNVLHGIDGPAVALKPDVEHAPSVSSVVEEDAEIVDTASK 121
D 3315 PSLSDVATVNVNVDINDNTFVFQSDYITTYISDAVLEQSV-ITVMADDA----DPSN 3369
QY 122 PGLDERLOKHDISESVNFDEETDGISQACLEPNSASSONSTDTGSGSATAAQPADN- 180
D 3370 SHIYSIIDGNGSSFTIDPVGRGVKTKILDRFT-----ISGYTLVVOASDNG 3418
QY 181 -----LIGDIDDDLE-----FVSPAPQGVTVRCRIIRKRGMDRGLFP 219
D 3419 SPRAVNTTIVNIDVDVNDNAVPSRGNYSVITQENKPVGFSLQVLVTDDESSHNG--P 3476

```

```

QY 220 TYVMYLEKENOKIF-----LLAARRKRSKTANYLISIDPYD----- 257
D 3477 PFFETITGNDKAFEVNPOGVLLTSSAIRKEDHLYLVKYADNGKPOLLSLTYYIDR 3536
QY 258 -----LSREGESYVGKLSNLMGTKFYYDR-----GICPMKG 290
D 3537 VIESIYPAIPLPLEIFITSSGEYSGVCKIHATDQVYDYLITSLDPQMDNLFSSVS 3596
QY 291 RGLVGAHTQOELALAIYETNNVLGFKPRKMSYIIPGMTLNHKQIPYQONNHDSLSRW 350
D 3597 TGGKLIHAKKLDIG--OYLLNVSVTDG--KFTTVADIVHIRQVT--QEMLNH-TIAIRF 3649
QY 351 QNRMENTLV-----ELHNKAPVNSDTSQY-----VLNPRGVTAQSVK 389
D 3650 ANLTPEEFVGDYMRNFORALRNILGYRRNDIQVLSQSSPPHPLDVLFEKPGSAQIS 3709
QY 390 NFQIVHK-----NDPDYIVNQFGRVADVDFTLDYVNPICAVQAFGIGLS-----S 434
D 3710 TKQLHKNSSVTDIEII-----GVRILNVFQKICA-----GLDCPMKFCDEKVS 3755
QY 435 FDKRI-----QILRKQELCELRHSHSASLVHRTAC-----QRWGVPMR 475
D 3756 VDESVMSTHSTARLSFTVPRH--HRAAVCLCKEGRCRPVHHCEDDPCPEGSECVSDPWE 3813
QY 476 Q 476
D 3814 E 3814

```

Search completed: July 16, 2002, 16:15:22
 Job time: 232 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 16, 2002, 16:10:21 ; Search time 30.7 Seconds
(without alignments)
2766.792 Million cell updates/sec

Title: US-09-782-390-1
Perfect score: 2568
Sequence: 1 MEASRCRLSPSGDSVFHEEM.....HPWROLPOSLVGPDLXLM 491

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp._archaea:*
2: sp._bacteria:*
3: sp._fungi:*
4: sp._human:*
5: sp._invertebrate:*
6: sp._mammal:*
7: sp._mhc:*
8: sp._organelle:*
9: sp._phage:*
10: sp._plant:*
11: sp._rodent:*
12: sp._virus:*
13: sp._vertebrate:*
14: sp._unclassified:*
15: sp._virus:*
16: sp._bacteria:*
17: sp._archaeop:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1089.5	42.4	420	4 Q9H4D2	Q9H4D2 homo sapien
2	994	38.7	358	13 Q9YH17	Q9YH17 gallus gall
3	916	35.7	543	11 Q9Z273	Q9Z273 mus musculu
4	882.5	34.4	443	5 Q9W2U5	Q9W2U5 drosophila
5	782.5	30.5	426	5 Q9UB22	Q9UB22 caenorhabdi
6	645.5	25.1	373	4 Q9GJ50	Q9GJ50 homo sapien
7	448	17.4	407	10 Q82257	Q82257 arabidopsis
8	427.5	16.6	455	10 Q9ZP59	Q9ZP59 arabidopsis
9	423.5	16.5	445	10 Q9ZP59	Q9ZP59 arabidopsis
10	405.5	15.8	462	10 Q9ZP78	Q9ZP78 oryza sativ
11	399.5	15.6	380	10 Q9SC01	Q9SC01 arabidopsis
12	397.5	15.5	415	10 Q9XIF9	Q9XIF9 arabidopsis
13	397.5	15.5	428	10 Q944S3	Q944S3 arabidopsis
14	396	15.4	379	10 Q93Y18	Q93Y18 arabidopsis
15	392	15.3	388	10 Q9C6B4	Q9C6B4 arabidopsis
16	391.5	15.2	411	10 Q9M3H9	Q9M3H9 cicer ariet

17	378.5	14.7	579	10 Q9MAG9	Q9MAG9 arabidopsis
18	362	14.1	428	10 Q9SXI3	Q9SXI3 lema pauci
19	359	14.0	368	10 Q94DR9	Q94DR9 oryza sativ
20	332	12.9	386	10 Q9ZPW1	Q9ZPW1 arabidopsis
21	242	9.4	397	10 Q9S9M6	Q9S9M6 arabidopsis
22	223	8.7	345	5 Q9N995	Q9N995 leishmania
23	214	8.3	1478	5 Q9YB18	Q9YB18 drosophila
24	209	8.1	53	6 Q9X833	Q9X833 equus caball
25	193.5	7.5	256	11 Q9Z2C2	Q9Z2C2 mus musculu
26	193.5	7.5	1547	11 Q9JIT5	Q9JIT5 mus musculu
27	190.5	7.4	1544	4 Q9NRJ4	Q9NRJ4 homo sapien
28	188	7.3	329	5 Q9BL53	Q9BL53 caenorhabdi
29	131	5.1	545	3 Q9HEK1	Q9HEK1 neurospora
30	127.5	5.0	839	10 Q9ZW32	Q9ZW32 arabidopsis
31	124	4.8	383	10 Q9M009	Q9M009 arabidopsis
32	115.5	4.5	949	5 Q95ZL5	Q95ZL5 caenorhabdi
33	115.5	4.5	1112	5 Q95ZL4	Q95ZL4 caenorhabdi
34	114	4.4	1325	4 Q9COD1	Q9COD1 homo sapien
35	114	4.4	2061	4 Q9BYW2	Q9BYW2 homo sapien
36	112	4.4	472	16 Q981V5	Q981V5 rhizobium 1
37	112	4.4	1045	5 Q9Y171	Q9Y171 drosophila
38	111.5	4.3	618	10 Q9MA81	Q9MA81 arabidopsis
39	111.5	4.3	1055	10 Q9MA51	Q9MA51 arabidopsis
40	111	4.3	977	4 Q95300	Q95300 homo sapien
41	111	4.3	977	4 Q9H476	Q9H476 homo sapien
42	111	4.3	1092	2 Q70022	Q70022 streptococ
43	110.5	4.3	673	5 Q9W137	Q9W137 drosophila
44	110.5	4.3	1465	10 Q9ZVY9	Q9ZVY9 arabidopsis
45	110	4.3	942	10 Q9W9L7	Q9W9L7 arabidopsis

ALIGNMENTS

RESULT	1	ALIGNMENTS
ID	Q9H4D2	PRELIMINARY; PRT; 420 AA.
AC	Q9H4D2	
DT	01-MAR-2001 (TREMBLrel. 16, Created)	
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)	
DT	01-OCT-2001 (TREMBLrel. 18, Last annotation update)	
DE	TUBBY (MOUSE) HOMOLOG (FRAGMENT).	
OS	TUB.	
GN	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	
OX	NCBI_TaxID=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Blood;	
RA	Brueckmann T., Winterpacht A., Hankeln T., Schmidt E.R., Zabel B.U.;	
RT	"Human PAC clone RPICP704G1312 (pbl 12G13), sequenced in DBCP Project	
RT	(Comparative Sequencing of a 1 Mb Region in Man (Chromosome 11p15) and	
RT	Mouse (Chromosome 7))."	
RL	Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.	
DR	EMBL: AI277662; CAC14586.1; -	
DR	HSSP: P50586; 1C87.	
DR	InterPro: IPR000007; Tubby.	
DR	Pfam: PF01167; TUB; 1.	
DR	PROSITE: PS01200; TUB_1; 1.	
FT	NON_TER	
SO	SEQUENCE 420 AA; 45642 MW; 0D195A55E39AD044 CRC64;	
Query Match	42.4%; Score 1089.5; DB 4; Length 420;	
Best Local Similarity	57.1%; Pred. No. 9,5e-78;	
Matches 213; Conservative	65; Mismatches 74; Indels 21; Gaps 8;	
QY	85 GIDGPAAVLK-PPEVTPAPSVSSVVEED-AENTVTASRPG-----LQELKQHDISE 135	
DB	48 GTSGPALAEKSEAGPQVILTVGGSDHAQDAGETFAAGGSPGQDRLRATVQRKGISS 107	
QY	136 SVNFDE----ETDGISQASCLE---RPNSASSQNSITDTGSSATA-AQPAINLLGDIDD 187	

```

Db 108 SMSFDEDEDEESSSSQSLNSNTRPSSATSRKSVESAASAPSTAPQPDV--VEQOD 164
Qy 188 LEDFYSPAGOVYVRCITLRDKRGMDGLFPTYMYLEKEENOKIFLLAARRKKSKSTA 247
Db 165 LEEFLYLRAPAGCITTKCITTRDKKMDGMPTTFLLHLDREDGKVFLLAARRKKSKSTS 224
Qy 248 NYLISIDEPVDSREGESYVGKLRSLNMGTKFTFYDRCIGCPMKRGVLGAHT-ROELAI 306
Db 225 NYLISVDPDLTDRGDSYIGKLRSLNMGTKFTFYDNGVNPQKASSSTLSECTLQELAAV 284
Qy 307 SYETVNLGFKPRKMSVITIPGTLNHHKOIPYQPONNHDSLSRWONKRMENLVELHAKAP 366
Db 285 CYETVNLGFKPRKMSVITIPGTLNHHKOIPYQPONNHDSLSRWONKRMENLVELHAKAP 344
Qy 367 VMSNDTQSYVLNFGRTVQASVKNFOIYHKNDPDYIWMQGRVADVTLDYNYPLCAVQ 426
Db 345 VMNDTQSYVLNFGRTVQASVKNFOIYHKNDPDYIWMQGRVADVTLDYNYPLCALQ 404
Qy 427 AFGIGLSEFDRKI 439
Db 405 AFAIALSSEDSKL 417

```

RESULT 2

```

Qy 09YH17 PRELIMINARY; PRT; 358 AA.
AC 09YH17:
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE TUBBY-LIKE PROTEIN.
GN TULP1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21376134; PubMed=11483369;
RA Heikenwaelder M.F., Koritschoner N.P., Pajer P., Chaboisier M.C.,
RA Kurz S.M., Brigel K.J., Bartunek P., Zenke M.;
RT "Molecular cloning, expression and regulation of the avian tubby-like
RT protein 1 (tulpi) gene."
RL Gene 273:131-139(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Koritschoner N.P., Zenke M.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Heikenwaelder M., Zenke M.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U92545; AAD09250.2; -.
DR HSSP; P50586; 1C8Z.
DR InterPro; IPR000007; Tubby.
DR Pfam; PF01167; Tub. 2.
DR PROSITE; PS01200; TUB_1; UNKNOWN_1.
SQ SEQUENCE 358 AA; 40123 MW; 9FF308649879118B CRC64;

```

Query Match 38.7%; Score 994; DB 13; Length 358;
 Best Local Similarity 54.1%; Pred. No. 2,6e-70;
 Matches 196; Conservative 58; Mismatches 72; Indels 36; Gaps 7;

```

Qy 95 PDEVHAPSVSSVEEDAEVTATSKP-----GLQERLQKHDSISV 137
Db 13 PSLCLPPATGSDDEDDSD---DASTKPIRSKKNPASLFTQGDGPPREKTRKKAPKTA 68
Qy 138 NDEETDGIQSACLERNSASSONSSTGTSGSATAQPADNLGLDDLEDEYVSAP 197
Db 69 ESEET-----LE--TQOKNSNKKGKRSKKKEERPLSPVI-EVDNLEKVELSPAP 117

```

```

Qy 198 QGYTVRCRITRDKRGMDGLFPTYMYLEKEENOKIFLLAARRKKSKTANYLISIDPD 257
Db 118 QGYTICRVRTDRKRGMDRGILPTHYHLND--KKFLLAGRRKKSKTSNLSIDPD 175
Qy 258 LSRGEGSYVGKLRSLNMGTKFTFYDRCIGCPMKRGVLGAHTROELAISETNVLGFKG 317
Db 176 LSRGEGSYVGKLRSLNMGTKFTFYDRCIGCPMKRGVLGAHTROELAISETNVLGFKG 233
Qy 318 PRKMSVITIPGTLNHHKOIPYQPONNHDSLSRWONKRMENLVELHAKAPVMSNDTQSYVL 377
Db 234 PRKMSVITIPGTLNHHKOIPYQPONNHDSLSRWONKRMENLVELHAKAPVMSNDTQSYVL 293
Qy 378 NFGRTVQASVKNFOIYHKNDPDYIWMQGRVADVTLDYNYPLCAVQAFGIGLSEFDRKI 437
Db 294 NFGRTVQASVKNFOIYHKNDPDYIWMQGRVADVTLDYNYPLCAVQAFGIGLSEFDRKI 353
Qy 438 RI 439
Db 354 KL 355

```

RESULT 3

```

Qy 09Z273 PRELIMINARY; PRT; 543 AA.
AC 09Z273:
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE TUBBY LIKE PROTEIN 1.
GN TULP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=BALB/CJ;
RA Nagert J.K., Nishina P.M., Fitch D., McGinnis N., Basson M., Yan G.,
RA Cardon L., Shiva N., Duyao M., Ikeda A., McGinnis A., North M.A.;
RT "Molecular evolution of the tubby gene family."
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=BALB/CJ;
RA Ikeda S., Sorokina I., Nagert J.K., North M.A., Nishina P.M.;
RT "Apoptotic photoreceptor cell death in tubby mice and the localization
RT of tubby gene family members in the retina."
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF085681; AAD13757.1; -.
DR EMBL; AF105711; AAD38451.1; -.
DR HSSP; P50586; 1C8Z.
DR MGD; MGT:109571; Tulp1.
DR InterPro; IPR000007; Tubby.
DR Pfam; PF01167; Tub. 1.
DR PROSITE; PS01200; TUB_1; UNKNOWN_1.
DR PROSITE; PS01201; TUB_2; 1.
SQ SEQUENCE 543 AA; 60335 MW; 0F0D9EAD2A2731E7 CRC64;

```

Query Match 35.7%; Score 916; DB 11; Length 543;
 Best Local Similarity 45.4%; Pred. No. 6,8e-64;
 Matches 199; Conservative 73; Mismatches 112; Indels 54; Gaps 10;

```

Qy 22 KMQQAKIDYRLLEKRRKKRLEPPMVQPNPEA-----RLRRKKPRAS 65
Db 137 KKEKSSLPKPKAKERKKAKAGPBGDVSPDPAPKPLRTKKKEVEGCTKLRAKKKKP 196
Qy 66 DE--QMPVLNCHPHSNVILHIGDPAVALKPEVHAPSVSSVVED--AEVTATSKP 122
Db 197 GELTDKP-----AGSPALARK---EFPAAMPLVGGGAAGVGKKKGP 237
Qy 123 GLQERLQKHDSISVNFDEETDGIQSACLERNSASSONSSTGTSGSATAQPADNL 182

```


RESULT	6	
096J50		
ID	096J50	PRELIMINARY; PRT; 373 AA.
AC	096J50;	
DT	01-DEC-2001 (TREMBLrel. 19, Created)	
DT	01-DEC-2001 (TREMBLrel. 19, last sequence update)	
DT	01-DEC-2001 (TREMBLrel. 19, last annotation update)	
DE	MOUSE TUBBY HOMOLOGUE (FRAGMENT).	
GN	TUB.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
CC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
OX	NCBI_TaxID=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A.	

RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhagen G.P., Preuss D., Nieman W.C., White O., Eissen J.A.,
RA Salberger S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant *Arabidopsis*
RA thaliana";
NL Nature 402:761-768(1999).

RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, COLUMBIA;
 RA Lin X;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC005309; AAC63644.1; -
 DR HSSP: P50586; 1C82.
 DR InterPro: IPR001810; F-box.
 DR InterPro: IPR000007; Tubby.
 DR Pfam: PF00646; F-box; 1.
 DR Pfam: PF01167; Tub; 1.
 SQ SEQUENCE 407 AA; 45382 MW; 45AA55907F09DE86 CRC64;

Query Match 17.4%; Score 448; DB 10; Length 407;
 Best Local Similarity 36.5%; Pred. No. 3.8e-27;
 Matches 112; Conservative 45; Mismatches 70; Indels 80; Gaps 10;

QY 195 PAPQGVYVRCRIIRKRGMDGLPPTYMYL-----EKEENQKIFLLAARRKKSKTANY 249
 DB 116 PGPRLSVCCYIMRNRNO-----TYLLYLGLNOAASNDGKFLAARFRPTCTDY 169
 QY 250 LLSIDPVDLSRGESEYVGLKRSNLMGTFEYD-----RGICPM 288
 DB 170 IISLNDVSRGNSNYIGKRSNFLTGFETYYDAQPTMPGTQVTRTSSRLSLKQVSPR 229
 QY 289 KRGVLGAHTRQELAISYETNVLGFKGPRKM-----SVITGMLNHQIYYOP 339
 DB 230 IFSGNPVAH-----ISYELNVLGSRGPRMOCVMDAIPASAVEPGT-----APQT 277
 QY 340 QNNHSL-----LSRMONRTM-----ENLVELHKKAPWNSDTQSYLVNF 379
 DB 278 ELVHSLVSPSPSPFRSSIRAESLPSGSPSSAAKQEGILLKAKAPRNHQLQCLNF 337
 QY 380 RGVHTQASVKNFOIV--HKNDP-----DYVMQFGVADVFTLDYNYPLCAVOAFGIGL 432
 DB 338 NGKVVTVASAKNQLVAAPENGPAGEHENVLIQFGKVGKDYFMDYQYPIAFAQFTICL 397
 QY 433 SSFDRKI 439
 DB 398 SSFDRKI 404

RESULT 8
 Q9ZP59 PRELIMINARY; PRT; 455 AA.
 AC Q9ZP59;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
 DE F22K20.1 PROTEIN (TUB FAMILY PROTEIN, PUTATIVE).
 GN F22K20.1 OR F7012.7.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, COLUMBIA;
 RA Federspiel N.A., Palm C.J., Conway A.B., Kurtz D.B., Conway A.R.,
 RA Au M., Araujo R., Buehler E., Dewar K., Feng C., Kim C., Li Y.,
 RA Oji O., Osborne B.I., Shinn P., Sun H., Toriumi M., Vysotskaia V.S.,
 RA Yu G., Becker J., Theologis A., Davis R.W.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 [2]
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, COLUMBIA;
 RX MEDLINE=21016719; PubMed=1130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,

RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzer L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.-J., Koo H.L., Kremetskaia I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Luos J.S., Malt R., Marshall A.,
 RA Miltischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 thaliana";
 RL Nature 408:816-820(2000).
 DR EMBL: AC002291; AAC00626.1; -
 DR EMBL: AC079283; AAG51146.1; -
 DR HSSP: P50586; 1C82.
 DR InterPro: IPR001810; F-box.
 DR InterPro: IPR000007; Tubby.
 DR Pfam: PF00646; F-box; 1.
 DR Pfam: PF01167; Tub; 1.
 DR PROSITE: PS01201; TUB_2; 1.
 SQ SEQUENCE 455 AA; 51194 MW; 53072DE655960C3E CRC64;

Query Match 16.6%; Score 427.5; DB 10; Length 455;
 Best Local Similarity 31.8%; Pred. No. 1.8e-25;
 Matches 109; Conservative 51; Mismatches 74; Indels 109; Gaps 11;

QY 195 PAPQGVYVRCRIIRKRGMDGLPPTYMYLKE-----ENQKIFLLAARRKKSKTANY 249
 DB 121 PGPRLATMOCFLIKRKSNNL-----TYHLYLCLSPALLVENGK-FLLSAKRIRRTTYTEY 173
 QY 250 LLSIDPVDLSRGESEYVGLKRSNLMGTFEYD-----RGICPMKRGVLGAHT 299
 DB 174 VISMADTTSRNSNYIGKRSNFLTGFETYYDQPTMPGTQVTRTSSRLSLKQVSPR 230
 QY 300 RQ-----ELAISYETNVLGFKGPRKM-----SVITPG 327
 DB 231 KRVSPKVPSSGYKINQVSEYLVNLTGRGPRNHCAMNSIPASLAEGCTVGPQDIIVPR 290
 QY 328 MTLNH-----KQIPQPNHDS-----LSRWKRTMENVEL 361
 DB 291 SLIDESFRSITSSSRKITYDSSNDFSSARFSDILGPLEDQEVVLEBGRKNSPVL-L 349
 QY 362 HNKAPWNSDTQSYVNFGRVTVASVKNFOIVHKND----- 398
 DB 350 KNRPRMNHQCLQCLNFGKRVTVASVKNFOIVANQPOPOPOPOPLTQPOPSGQT 409
 QY 399 --PDYVMQFGVADVFTLDYNYPLCAVOAFGIGLSFDRKI 439
 DB 410 DGPDKIILQFGKVKGMFTMDYFPLSAFAICLSSFDTKL 452

RESULT 9
 Q9FRH7 PRELIMINARY; PRT; 445 AA.
 AC Q9FRH7;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
 DE HYPOTHETICAL 50.0 KDA PROTEIN.
 GN F4F7.33.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, COLUMBIA;
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,

QY 197 POGATVACRIIRIKRGMGRGJEPYVYL-----EKEKNIPIILAAARKKSTANYLIS 25
Db 100 PRDVLQCFIKRNRNTQ-----SYHLGLGITSLDNGK-FLNAASKIKRITCTDIYIS 15
QY 253 IDPVYDISRSGESVYAKLSSNLMGRKFTVYDRG-----ICPMGRGLVG 295
Db 153 LRRDDISKSNMAYLGRMNSNGLGRFVYFGSOTGAARKOKSRSSNFTKVSFVPOGSRP 212

RESULT	14
093VI8	
ID	093VI8
AC	093VI8;
DT	01-DEC-2001 (TREMBLrel. 19, Created)
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE	PTATIVE-TUB FAMILY PROTEIN.
GN	F12M16.22.
OS	Arabidopsis thaliana (mouse-ear cress).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC	Brassicales; Brassicaceae; Arabidopsids.
NCBI_Taxid=3702;	
PRT;	379 AA.

[1]
 RN SEQUENCE FROM N.A.
 RP Yamada K., Banh J., Banno F., Chang E., Dale J.M., Goldsmith A.D.,
 RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,
 RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
 RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Kim C., Koeseema E., Lam B., Lin J., Meyers M.C., Miranda M.,
 RA Narasaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
 RA Shinn P., Southwick A., Tracy S.E., Shinozaki K., Davis R.W.,
 RA Ecker J.R., Theologis A., Theologis A., Theologis A.,
 RT "Full Length cDNA of gene F12M16.22 (GI:7769867)."
 RN Submitted (Oct-2001) to the EMBL/Genbank/DBJ databases.
 RL
 RP SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
 RA Goldsmith A.D., Lee J.M., Quach H.L., Tang C., Toriumi M., Yu G.,
 RA Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J.,
 RA Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Koeseema E.,
 RA Lam B., Lin J., Meyers M.C., Miranda M., Narasaka M., Nguyen M.,
 RA Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A.,
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.,
 RT "Full Length cDNA of gene F12M16.22 (GI:7769867)."
 RN Submitted (Apr-2001) to the EMBL/Genbank/DBJ databases.
 RL EMBL: AY059088; AAL15194.1;
 DR EMBL: AF370146; AAK43961.1;
 SO SEQUENCE 379 AA; 42178 MW; B31C395DE9E1164 CRC64;

Query Match 15.4%, Score 396, DB 10, Length 379;
 Best Local Similarity 29.8%; Pred. No. 4.3e-23;
 Matches 119; Conservative 57; Mismatches 115; Indels 108; Gaps 12;

QY 124 LGERLKHHSFVNDE-ETDGISOSACLERPNSASSONSDTSGSATAAPADNLL 182
 DB 3 LRSRLSRISNSFRHGETTAPESISTPPSN-----MASSSSAMLPBL 52
 QY 183 GDI-----DLEDFVYS-----PAPOG 199
 DB 53 GELIRVEETEDRMPQRDVYTCACVSKKREITHDFARSSINSKITEPSCLEKLPGRD 112
 QY 200 VYVRCIITDKRGMDRGLFPTTYMVL--EKENOKIFLLAARKKKSKTANYLISDPV 256
 DB 113 FSNOCILIKRKK-----STFYLYLALTPSPFDKGLFLAARRRTGAYEYIISLAD 166
 QY 257 DLSREGESVYGLKRLSMLGKTFYVD-----RGICPMRGVLGAHTRO----- 301
 DB 167 DTSOGSNAYVGLKRLSDFLTSTFTYDSDPPHNGAKPSNCKASRRRA-SKQISPOVPAGNF 225
 QY 302 ELAISTYETNVLGFGPKRKMSYIIPGMLNKHQIPYQDNNHDSLLSRWQ----- 351
 DB 226 EGVHYSYKNNLKLKSGPRRM-----VSTLRCPSPSPSSSSAGLSQKPCDVTIMK 277
 QY 352 --NRTMENLVEILHKKAPVWNSDTSQSYVLNFRGRVYQASVKNQIV-----HKNDPD 400
 DB 278 KPNKGGSSILTLKNAKPRWHEHLQCMCNFNGRVVAKNQLAVAYDOSOPSGDDE 337
 QY 401 YIVMOGRVADVFETLDYNYPLCAVOAFGIGLSEFDRK1 439
 DB 338 TVLDQFGKVGDDTFTMDYRQPLSAFOAFALCITSPGTL 376

RESULT 15
 O9C6B4 PRELIMINARY: PRT: 388 AA.
 AC O9C6B4:
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE HYPOTHETICAL 43.3 KDA PROTEIN.
 GN F8G22.1.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CY, COLUMBIA;
 RX MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federpspiel N.A., Kaul S.,
 RA White O., Alonso J., Altafi H., Alarajo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Cressy T.H., Dewar K.,
 RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hulzer L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Malet R., Marziani A.,
 RA Millscher J., Miranda M., Nguyen M., Nieman W.C., Osborne B.I.,
 RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 RA Uterback T., Van Aken S., Vayenberg M., Vysotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.,
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 thaliana."
 RT Nature 408:816-820(2000).
 RL EMBL: AC079677; AAG52638.1;
 DR HSSP: P50586; 1C8Z.
 DR InterPro: IPR001810; F-box.
 DR InterPro: IPR000007; Tubby.
 DR Pfam: PF00646; F-box; 1.
 DR Pfam: PF01167; Tub; 1.
 DR PROSITE: PS01200; TUB_1; UNKNOWN_1.
 KW Hypothetical protein.
 SO SEQUENCE 388 AA; 43340 MW; 00EA4D1BEF90BA30 CRC64;

Query Match 15.3%, Score 392, DB 10, Length 388;
 Best Local Similarity 29.9%; Pred. No. 9.2e-23;
 Matches 112; Conservative 64; Mismatches 105; Indels 94; Gaps 13;

QY 156 NSASSNSTDTGTSGSAQAQPADNLLGDIIDLED----- 190
 DB 27 SSVSSLSLSTEGMLQSIWVDLPPELLDITQRIESESQSLMPGRDVAACASVCKSREMT 86
 QY 191 -----FVSPAPQGVTVRCRIIRK-RGMDR---GLPFTYMYLEKEE 229
 DB 87 KEVYKVELSGILTFPISLQPGRRAPICQFKIKREATIYLLYLGLSRA-----LSGDK 142
 QY 230 NOKIFLLAARRKKSKTANYLISIDPVDLSREGESYVGLKRLSMLGKTFYVD----- 282
 DB 143 SK--LLLSAKRVRRATGAEFVYLSGNDFSRSSNNTIGKRLSNFLGKFFVYENQPPPN 200
 QY 283 RGICP-WKGRGLVAAATRQELAAISTETNVLGFGPKRKMSYI--IP-----GMTLN 331
 DB 201 RKLPPSQMVSPWSSSSSYNIAISILVELNVLTRGRRMQCLMHSIPISAIDGKIOS 260
 QY 332 HKQIPYQDNNHDSLLSRWQ--NRTMENLVE---LHKKAPVWNSDTSQSYVLNFRGVYQTA 386
 DB 261 PFEETNGCKKKKKRKLMPGSCNGLGSEVITKEPLTKKSPRWHEHLQCMCNLKNKGRVTA 320
 QY 387 SVKNFQIV-----HKNDPDIYVMOGRVADVFETLDYNYPLCAVOAFGIGLSS 434
 DB 321 SVKNFQIVAAAEAGKNNMIPEDQDVIIIOFGIKGIDFTMDYRVPISAFQAFALCIS 380
 QY 435 FDKRIQTLRMQELCE 449
 DB 381 FDKR-----PVCE 388

Search completed: July 16, 2002, 16:14:31
 Job time: 250 sec